

# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

BEST AVAILABLE COPY

## STAFF USE ONLY

Date completed: <u>01-14-03</u>	Search Site	Vendors
Searcher: <u>Beverly C. 20 2528</u>	<input type="checkbox"/> STIC	<input type="checkbox"/> IG
Terminal time: _____	<input type="checkbox"/> CM-1	<input type="checkbox"/> STN
Elapsed time: _____	<input type="checkbox"/> Pre-S	<input type="checkbox"/> Dialog
CPU time: _____	Type of Search	<input type="checkbox"/> APS
Total time: <u>25</u>	<input type="checkbox"/> N.A. Sequence	<input type="checkbox"/> Geninfo
Number of Searches: _____	<input type="checkbox"/> A.A. Sequence	<input type="checkbox"/> SDC
Number of Databases: <u>1</u>	<input type="checkbox"/> Structure	<input type="checkbox"/> DARC/Questel
	<input type="checkbox"/> Bibliographic	<input checked="" type="checkbox"/> Other <u>CGN</u>



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 13, 2004, 15:53:18 ; Search time 744 Seconds  
(without alignments)  
1044.735 Million cell updates/sec

Title: US-09-719-737-18

Perfect score: 19  
Sequence: 1 ctgggccatcagtcctcg 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_da:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_cm:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_stb:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vt:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_ph:\*

24: em\_pi:\*

25: em\_ro:\*

26: em\_sts:\*

27: em\_un:\*

28: em\_vt:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rtd:\*

36: em\_htg\_man:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	100.0	19	6 AX008665	AX008665 Sequence
2	19	100.0	19	6 AX008666	AX008666 Sequence
3	19	100.0	19	6 AX671175	AX671175 Sequence
4	19	100.0	260	9 AF224495	AF224495 Homo sapi
5	19	100.0	348	9 AF262299	AF262299 Homo sapi
6	19	100.0	406	9 AF262301	AF262301 Homo sapi
7	19	100.0	410	9 AF262302	AF262302 Homo sapi
8	19	100.0	1065	6 AR164119	AR164119 Sequence
9	19	100.0	1065	6 AX030929	AX030929 Sequence
10	19	100.0	1065	6 BD128569	BD128569 Sequence
11	19	100.0	1068	6 AX280851	AX280851 Sequence
12	19	100.0	1068	6 AX323054	AX323054 Sequence
13	19	100.0	1068	9 AB023887	AB023887 Homo sapi
14	19	100.0	1068	9 AF026535	AF026535 Homo sapi
15	19	100.0	1068	9 AY221092	AY221092 Homo sapi
16	19	100.0	1071	6 AR278855	AR278855 Sequence
17	19	100.0	1071	6 BD015209	BD015209 Mammalian
18	19	100.0	1116	6 AR300123	AR300123 Sequence
19	19	100.0	1116	6 BD082062	BD082062 G-protein
20	19	100.0	1193	6 AR300122	AR300122 Sequence
21	19	100.0	1193	6 BD082061	BD082061 G-protein
22	19	100.0	1201	6 AR270522	AR270522 Sequence
23	19	100.0	1201	6 AX548778	AX548778 Sequence
24	19	100.0	1201	9 HSU28694	HSU28694 Human eosin
25	19	100.0	1310	9 AF262300	AF262300 Homo sapi
26	19	100.0	1689	6 AR300121	AR300121 Sequence
27	19	100.0	1689	6 BD082060	BD082060 G-protein
28	19	100.0	1689	9 HSU49727	HSU49727 Human C-C
29	19	100.0	1717	6 AX34894	AX34894 Sequence
30	19	100.0	1717	9 HSU51241	HSU51241 Human eosin
31	19	100.0	1915	6 BD006761	BD006761 Chemokine
32	19	100.0	1915	6 BD017703	BD017703 Chemokine
33	19	100.0	2030	9 BC033514	BC033514 Homo sapi
34	19	100.0	5791	6 AX705064	AX705064 Sequence
35	19	100.0	5791	9 AF247361	AF247361 Homo sapi
36	19	100.0	7010	9 AF224496S2	AF224497 Homo sapi
37	19	100.0	177334	9 AC138069	AC138069 Homo sapi
38	19	100.0	197279	9 AC104439	AC104439 Homo sapi
39	19	100.0	220965	2 HSA312688	AJ312688 Homo sapi
40	17.4	91.6	1508	10 BC030031	BC030031 Mus muscu
41	17.4	91.6	1511	10 AF248636	AF248636 Mus muscu
42	17.4	91.6	1516	6 AX180365	AX180365 Sequence
43	17.4	91.6	2445	10 AF248635	AF248635 Mus muscu
44	17.4	91.6	110000	2 AC108356_1	Continuation (2 of
45	17.4	91.6	135394	10 AL672140	AL672140 Mouse DNA

## ALIGNMENTS

RESULT 1  
AX008665 19 bp DNA linear PAT 06-SEP-2000  
LOCUS AX008665  
DEFINITION Sequence 18 from Patent WO9966037.  
ACCESSION AX008665  
VERSION AX008665.1 GI:9996189  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1  
Renzl, P.  
Antisense oligonucleotides for treating or preventing atopic  
diseases and neoplastic cell proliferation  
Patent: WO 9966037-A 18 23-DEC-1999;  
JOURNAL

FEATURES  
source RENZI PAOLO (CA) ; RECH EXPERTISES ET DEV MEDICAU (CA)  
Location/Qualifiers  
1. .19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Antisense oligonucleotide inhibiting the CCR3 human receptor"

BASE COUNT 2 a 6 c 6 g 5 t

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTGCTCTG 19  
|||||  
1 CTGGGCCATCAGTGCTCTG 19

Db 1 CTGGGCCATCAGTGCTCTG 19

RESULT 2  
AX008666/c  
LOCUS AX008666 19 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 19 from Patent WO966037.  
ACCESSION AX008666  
VERSION AX008666.1 GI:9996190  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Renzi, P.  
TITLE Antisense oligonucleotides for treating or preventing atopic diseases and neoplastic cell proliferation  
JOURNAL Patent: WO 9966037-A 19 23-DEC-1999;  
RENTZI PAOLO (CA) ; RECH EXPERTISES ET DEV MEDICAU (CA)  
FEATURES Location/Qualifiers  
source 1. .19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Sense oligonucleotide for the CCR3 human receptor"

BASE COUNT 5 a 6 c 6 g 2 t

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTGCTCTG 19  
|||||  
19 CTGGGCCATCAGTGCTCTG 1

Db 19 CTGGGCCATCAGTGCTCTG 1

RESULT 3  
AX671175  
LOCUS AX671175 19 bp DNA linear PAT 27-MAR-2003  
DEFINITION Sequence 15 from Patent WO03004511.  
ACCESSION AX671175  
VERSION AX671175.1 GI:29329631  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Renzi, P., Allam, M. and Allakhverdi, Z.  
TITLE Method for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
JOURNAL Patent: WO 03004511-A 15 16-JAN-2003;  
Topigen Pharmaceuticals Inc (CA)  
FEATURES Location/Qualifiers  
source 1. .19  
/organism="synthetic construct"

BASE COUNT 2 a 6 c 6 g 5 t

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTGCTCTG 19  
|||||  
1 CTGGGCCATCAGTGCTCTG 19

Db 1 CTGGGCCATCAGTGCTCTG 19

RESULT 4  
AF224495/c  
LOCUS AF224495 260 bp mRNA linear PRI 02-MAY-2001  
DEFINITION Homo sapiens CC chemokine receptor 3 (CCR3) mRNA, partial cds.  
ACCESSION AF224495  
VERSION AF224495.1 GI:13924481  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Scotet, E.J.  
TITLE CCR3 expression is associated with chromatin remodeling in Th2 cells  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 260)  
AUTHORS Scotet, E.J.  
TITLE Direct Submission  
JOURNAL Submitted (13-JUN-2000) Basel Institute for Immunology, 487 Grenzacherstrasse, Basel CH-4005, Switzerland  
FEATURES Location/Qualifiers  
source 1. .260  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_type="T cell"  
1. .>260  
/gene="CCR3"  
91. .>260  
/gene="CCR3"  
/codon\_start=1  
/product="CC chemokine receptor 3"  
/protein\_id="AAK49027.1"  
/db\_xref="GI:13924482"  
/translation="MTTSLDPTVETFGTTSYYDDVGLCEKADTRALMAQFVPPYSLV FTVGLGNVVVV"

BASE COUNT 64 a 57 c 66 g 73 t

ORIGIN

Query Match 100.0%; Score 19; DB 9; Length 260;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTGCTCTG 19  
|||||  
195 CTGGGCCATCAGTGCTCTG 177

Db 195 CTGGGCCATCAGTGCTCTG 177

RESULT 5  
AF262299/c  
LOCUS AF262299 348 bp mRNA linear PRI 26-JUN-2002  
DEFINITION Homo sapiens clone 1 CC chemokine receptor 3 (CCR3) mRNA, partial cds.  
ACCESSION AF262299  
VERSION AF262299.1 GI:19171640  
KEYWORDS  
SOURCE Homo sapiens (human)



ORGANISM Homo sapiens  
REFERENCE 2 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE 1 (bases 1 to 348)  
Vijh.S., Dayhoff,D.E., Wang,C.E., Imam,Z., Ehrenberg,P.K. and  
Michael,N.L.  
JOURNAL Transcription Regulation of Human Chemokine Receptor CCR3: Evidence  
MEDLINE for a Rare TATA-less Promoter Structure Conserved between  
22074933 Drosophila and Humans  
PUBMED 12079287  
2 (bases 1 to 348)  
Vijh.S., Dayhoff,D.E., Wang,C.E., Ehrenberg,P.K. and Michael,N.L.  
REFERENCE Direct Submission  
AUTHORS Submitted (23-MAR-2000) Dept. of Mol. Diag. and Pathogenesis,  
TITLE Walter Reed Army Institute of Research, 1600 E. Gude Drive,  
JOURNAL Rockville, MD 20850, USA  
FEATURES  
source 1. 348  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p21"  
/clone="1"  
/cell\_type="eosinophil"  
1. .>348  
/gene="CCR3"  
120. .>348  
/note="G-protein coupled receptor; principal eotaxin  
receptor expressed on eosinophils; CD4 Th2 lymphocytes,  
CD8 lymphocytes, microglia, dendritic cells, and  
monocytes"  
/codon\_start=1  
/product="CC chemokine receptor 3"  
/protein\_id="FAL85628.1"  
/db\_xref="GI:19171641"  
/translation="MTSLDVTETFGTSTYYDVGILCEKADTRALMAQVPPYLSLV  
FTVGLGNVVVWVILIKYRLRIMTNYLNL"  
BASE COUNT 94 a 79 c 81 g 94 t  
ORIGIN  
Query Match 100.0%; Score 19; DB 9; Length 348;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTGGGCGCATCAGTCTCTG 19  
Db 224 CTGGGCGCATCAGTCTCTG 206  
RESULT 6  
LOCUS AF262301/c 406 bp mRNA linear PRI 26-JUN-2002  
DEFINITION Homo sapiens clone 4 CC chemokine receptor 3 (CCR3) mRNA, partial  
c86.  
ACCESSION AF262301  
VERSION AF262301.1 GI:19171644  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 406)  
Vijh.S., Dayhoff,D.E., Wang,C.E., Imam,Z., Ehrenberg,P.K. and  
Michael,N.L.  
JOURNAL Transcription Regulation of Human Chemokine Receptor CCR3: Evidence  
MEDLINE for a Rare TATA-less Promoter Structure Conserved between  
22074933 Drosophila and Humans  
PUBMED 12079287  
2 (bases 1 to 406)  
Vijh.S., Dayhoff,D.E., Wang,C.E., Ehrenberg,P.K. and Michael,N.L.  
REFERENCE Direct Submission  
AUTHORS Submitted (23-MAR-2000) Dept. of Mol. Diag. and Pathogenesis,  
TITLE Walter Reed Army Institute of Research, 1600 E. Gude Drive,  
JOURNAL Rockville, MD 20850, USA  
FEATURES  
source 1. 410  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

PUBMED 12079287  
REFERENCE 2 (bases 1 to 406)  
AUTHORS Vijh.S., Dayhoff,D.E., Wang,C.E., Ehrenberg,P.K. and Michael,N.L.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAR-2000) Dept. of Mol. Diag. and Pathogenesis,  
Walter Reed Army Institute of Research, 1600 E. Gude Drive,  
Rockville, MD 20850, USA  
FEATURES  
source 1. 406  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p21"  
/clone="4"  
1. .>406  
/gene="CCR3"  
160. .>406  
/note="G-protein coupled receptor; principal eotaxin  
receptor expressed on eosinophils; CD4 Th2 lymphocytes,  
CD8 lymphocytes, microglia, dendritic cells, and  
monocytes"  
/codon\_start=1  
/product="CC chemokine receptor 3"  
/protein\_id="FAL85630.1"  
/db\_xref="GI:19171645"  
/translation="MTSLDVTETFGTSTYYDVGILCEKADTRALMAQVPPYLSLV  
FTVGLGNVVVWVILIKYRLRIMTNYLNLISDG"  
BASE COUNT 107 a 93 c 98 g 108 t  
ORIGIN  
Query Match 100.0%; Score 19; DB 9; Length 406;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTGGGCGCATCAGTCTCTG 19  
Db 264 CTGGGCGCATCAGTCTCTG 246  
RESULT 7  
LOCUS AF262302/c 410 bp mRNA linear PRI 26-JUN-2002  
DEFINITION Homo sapiens clone 5 CC chemokine receptor 3 (CCR3) mRNA, partial  
c86.  
ACCESSION AF262302  
VERSION AF262302.1 GI:19171646  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 410)  
Vijh.S., Dayhoff,D.E., Wang,C.E., Imam,Z., Ehrenberg,P.K. and  
Michael,N.L.  
JOURNAL Transcription Regulation of Human Chemokine Receptor CCR3: Evidence  
MEDLINE for a Rare TATA-less Promoter Structure Conserved between  
22074933 Drosophila and Humans  
PUBMED 12079287  
2 (bases 1 to 410)  
Vijh.S., Dayhoff,D.E., Wang,C.E., Ehrenberg,P.K. and Michael,N.L.  
REFERENCE Direct Submission  
AUTHORS Submitted (23-MAR-2000) Dept. of Mol. Diag. and Pathogenesis,  
TITLE Walter Reed Army Institute of Research, 1600 E. Gude Drive,  
JOURNAL Rockville, MD 20850, USA  
FEATURES  
source 1. 410  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

gene  
CDS  
/chromosome="3"  
/map="3p21"  
/clone="5"  
/cell\_type="primary eosinophil"  
1..>410  
/gene="CCR3"  
170..>410  
/gene="CCR3"  
/note="G-protein coupled receptor; principal eotaxin receptor expressed on eosinophils, CD4 Th2 lymphocytes, CD8 lymphocytes, microglia, dendritic cells, and monocytes"  
/codon\_start=1  
/product="CC chemokine receptor 3"  
/protein\_id="AA185631.1"  
/db\_xref="GI:19171647"  
/translation="MTTSLDVTETFTTSYDDVGLCEKADTRALMAQVPPPLSLVFTVGLGNVVMVILIKYRRRLRMITNITLNLALSD"

BASE COUNT 113 a 91 c 96 g 110 t  
ORIGIN

Query Match 100.0%; Score 19; DB 9; Length 410;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCTG 19  
Db 274 CTGGGCCATCAGTCTCTG 256

RESULT 8  
ARI64119/c ARI64119 1065 bp DNA linear PAT 17-OCT-2001  
LOCUS  
DEFINITION Sequence 2 from patent US 6271347.  
ACCESSION ARI64119  
VERSION ARI64119.1 GI:16235065  
KEYWORDS  
SOURCE Unknown.  
ORGANISM  
REFERENCE 1 (bases 1 to 1065)  
AUTHORS Daugherty,B.L., Demartino,J.A., Siciliano,S.J. and Springer,M.S.  
TITLE Eosinophil eotaxin receptor  
JOURNAL Patent: US 6271347-A 2 07-AUG-2001;  
FEATURES  
source 1..1065  
Location/Qualifiers  
BASE COUNT 230 a 289 c 242 g 304 t  
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCTG 19  
Db 105 CTGGGCCATCAGTCTCTG 87

RESULT 9  
AX030929/c AX030929 1065 bp DNA linear PAT 20-SEP-2000  
LOCUS  
DEFINITION Sequence 2 from Patent EP1012190.  
ACCESSION AX030929  
VERSION AX030929.1 GI:10278334  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1  
AUTHORS Daugherty,B.L., Demartino,J.A., Siciliano,S.J. and Springer,M.S.  
TITLE Eosinophil eotaxin receptor  
JOURNAL Patent: EP 1012190-A 2 28-JUN-2000;

FEATURES  
source  
MERCK & CO INC (US)  
Location/Qualifiers  
1..1065  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

BASE COUNT 230 a 289 c 242 g 304 t  
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCTG 19  
Db 105 CTGGGCCATCAGTCTCTG 87

RESULT 10  
BD128569/c BD128569 1065 bp DNA linear PAT 18-SEP-2002  
LOCUS  
DEFINITION Eosinophil eotaxin receptor.  
ACCESSION BD128569  
VERSION BD128569.1 GI:23223514  
KEYWORDS JP 2002503950-A/1.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1065)  
AUTHORS Daugherty,B.L., Demartino,J.A., Springer,M.S. and Siciliano,S.J.  
TITLE Eosinophil eotaxin receptor  
JOURNAL Patent: JP 2002503950-A 1 05-FEB-2002;  
COMMENT MERCK & CO INC  
OS Unidentified  
PN JP 2002503950-A/1  
PD 05-FEB-2002  
PF 24-APR-1997 JP 1997538970  
PR 26-APR-1996 US 08/640991,26-APR-1996 US 60/016158 PR  
17-JAN-1997 GB 9700894.0  
PI BRUCE L DAUGHERTY, JULIE A DEMARTINO, MARTIN S SPRINGER PI  
PC C07K14/705,C07K14/715,C12N15/12  
CC Strandedness: Single;  
CC Topology: Linear;  
CC Eosinophil eotaxin receptor  
FH Key  
FT source 1..1065  
Location/Qualifiers  
/organism="Unidentified".

FEATURES  
source  
Location/Qualifiers  
1..1065  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

BASE COUNT 230 a 289 c 242 g 304 t  
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCTG 19  
Db 105 CTGGGCCATCAGTCTCTG 87

RESULT 11  
AX280851/c AX280851 1068 bp DNA linear PAT 02-NOV-2001  
LOCUS  
DEFINITION Sequence 474 from Patent WO0177172.  
ACCESSION AX280851  
VERSION AX280851.1 GI:16608181  
KEYWORDS  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1  
 Lehmann-Brunsmann, K., Liaw, C.W. and Lin, I.L.  
 Non-endogenous, constitutively activated known g protein-coupled receptors  
 Patent: WO 0177172-A 474 18-OCT-2001;  
 Arena Pharmaceuticals, Inc. (US)

JOURNAL  
 Location/Qualifiers

FEATURES  
 source  
 1. 1068  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 232 a 288 c 244 g 304 t

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCATGCTCTG 19  
 |||||  
 105 CTGGGCATCATGCTCTG 87

Db 105 CTGGGCATCATGCTCTG 87

RESULT 12  
 AX323054/c  
 LOCUS AX323054 1068 bp DNA linear PAT 07-JAN-2002  
 DEFINITION Sequence 2 from Patent WO0192520.  
 ACCESSION AX323054  
 VERSION AX323054.1 GI:18093940  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1  
 Barnes, A.A., Fraser, N.J., O'Shaughnessy, C.T. and Wise, A.G.  
 Modified chemokine receptor ccr-3 and assay  
 Patent: WO 0192520-A 2 06-DEC-2001;  
 GLAXO GROUP LIMITED (GB)

JOURNAL  
 Location/Qualifiers

FEATURES  
 source  
 1. 1068  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1. 1068  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD20401.1"  
 /db\_xref="GI:18093941"  
 /translation="MTTSLDPTVEFTGTSYYDDVGLCEKADTRALMAQVPPPLYSIV  
 FTVGLGNVVVVMILIKYRRLIMTNYILNLAIISDLPLVLTLPFWIHYVGNWVVG  
 HGMCKLISGFYETELSEIFILLIDRYLAIVHAFALRARVFGVITSITWGL  
 AVLAALPEPFIYEETELSEIFELTCSALYEDTVSMRPHRLRMRTICLVPLVLAIS  
 YTGIIKITLACPSKKYKAIKRLIPIVAVAFIEMTPVNVAILSSYOSIIGNDCCERT  
 KHDLVNLYTVBVIAYSHCCNMPVIYAVGGRFKRLHFRHRLHMLGRIIPLPSE  
 KLERTSSVSPTAPELIVF"

BASE COUNT 231 a 289 c 243 g 305 t

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCATGCTCTG 19  
 |||||  
 105 CTGGGCATCATGCTCTG 87

Db 105 CTGGGCATCATGCTCTG 87

RESULT 13  
 AB023887/c

LOCUS AB023887 1068 bp DNA linear PRI 25-NOV-1999  
 DEFINITION Homo sapiens gene for b-chemokine receptor CCR3, complete cds.  
 ACCESSION AB023887  
 VERSION AB023887.1 GI:6467132  
 KEYWORDS b-chemokine receptor CCR3.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1068)  
 Kato, H., Tsuchiya, N., Izumi, S., Miyama, M., Nakajima, T., Kawasaki, H., Hirai, K. and Tokunaga, K.  
 New variations of human CC-chemokine receptors CCR3 and CCR4  
 Genes Immun. 1 (2), 97-104 (1999)  
 MEDLINE 21040311  
 PUBMED 11196669

REFERENCE 2 (bases 1 to 1068)  
 Kato, H., Tsuchiya, N. and Tokunaga, K.  
 Direct Submission  
 Submitted (19-FEB-1999) Hitooshi Kato, University of Tokyo,  
 Department of Human Genetics, 7-3-1 Hongo, Bunkyo-ku, Tokyo  
 113-0033, Japan (E-mail: kato@u-tokyo.ac.jp, Tel: 61-3-5841-3542,  
 fax: 61-3-5802-8619)

JOURNAL  
 Location/Qualifiers

FEATURES  
 source  
 1. 1068  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /map="3p21.2"  
 1. 1068  
 /codon\_start=1  
 /product="b-chemokine receptor CCR3"  
 /protein\_id="BA06964.1"  
 /db\_xref="GI:6467133"  
 /translation="MTTSLDPTVEFTGTSYYDDVGLCEKADTRALMAQVPPPLYSIV  
 FTVGLGNVVVVMILIKYRRLIMTNYILNLAIISDLPLVLTLPFWIHYVGNWVVG  
 HGMCKLISGFYETELSEIFILLIDRYLAIVHAFALRARVFGVITSITWGL  
 AVLAALPEPFIYEETELSEIFELTCSALYEDTVSMRPHRLRMRTICLVPLVLAIS  
 YTGIIKITLACPSKKYKAIKRLIPIVAVAFIEMTPVNVAILSSYOSIIGNDCCERT  
 KHDLVNLYTVBVIAYSHCCNMPVIYAVGGRFKRLHFRHRLHMLGRIIPLPSE  
 KLERTSSVSPTAPELIVF"  
 652  
 variation  
 /replace="t"

BASE COUNT 232 a 289 c 243 g 304 t

ORIGIN

Query Match 100.0%; Score 19; DB 9; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCATGCTCTG 19  
 |||||  
 105 CTGGGCATCATGCTCTG 87

Db 105 CTGGGCATCATGCTCTG 87

RESULT 14  
 AF026535/c  
 LOCUS AF026535 1068 bp mRNA linear PRI 02-NOV-1997  
 DEFINITION Homo sapiens chemokine receptor (CCR3) mRNA, complete cds.  
 ACCESSION AF026535  
 VERSION AF026535.1 GI:2582565  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1068)  
 Xiao, L., Weiss, S., Gari, S., Rudolph, D., Hodge, T. and Lai, R.  
 Partial resistance to infection by syncytium-inducing primary HIV-1  
 in exposed uninfected individuals homozygous for CCR3 32bp deletion  
 Unpublished  
 2 (bases 1 to 1068)

JOURNAL  
 Unpublished

AUTHORS Qari, S.  
TITLE Direct Submission  
JOURNAL Submitted (24-SEP-1997) Retrovirus Diseases Branch, Centers for  
Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA  
30333, USA

## FEATURES

source Location/Qualifiers

1..1068  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p21"  
1..1068  
/gene="CCR3"  
1..1068  
/gene="CCR3"  
/codon\_start=1  
/product="Chemokine receptor"  
/protein\_id="AAB82589.1"  
/db\_xref="GI:2582566"  
/translation="MTSLDVTETFGTTSYYDDVGLCEKADTRALMAQFVPLYSIV  
FTVGLGNVVVVMILIKYRLRIWNTNLYLNLASDLLFLVLPFWHYGNHVRG  
HGMCKLSGFHTGLYSIFILITIDRYLAIVHAFALRATVFGVTSITWGL  
AVLAALPEPIFYETEELPEETLCSALYPEDVYSWRHPTLRMTIFGLVPLVMAIC  
VTGIIKTLRCPSKKYKARLIFVIMAVFPIFTVPVVALILSSYQILGNDCCRS  
KHLDLVNLVTEVIAVSHCCANPVYAVGSRFRKYLRFHFRHLLMLHGRYIPLPSR  
KLRTSSVSPTAPELSIVF"

CDS

BASE COUNT 231 a 289 c 243 g 305 t  
ORIGIN

Query Match 100.0%; Score:19; DB 9; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 14;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
105 CTGGGCCATCAGTCTCTG 87

RESULT 15  
AY221092/c

LOCUS AY221092 1068 bp DNA linear PRI 11-APR-2003  
DEFINITION Homo sapiens chemokine receptor 3 gene, complete cds.  
ACCESSION AY221092  
VERSION AY221092.1 GI:29169288

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1068)  
Kopatz, S.A., Aronstam, R.S. and Sharma, S.V.

TITLE Isolation of complete coding sequence for chemokine (C-C motif)  
receptor 3 (CCR3)

UNPUBLISHED

2 (bases 1 to 1068)  
Kopatz, S.A., Aronstam, R.S. and Sharma, S.V.

REFERENCE 1 (bases 1 to 1068)  
Kopatz, S.A., Aronstam, R.S. and Sharma, S.V.

TITLE Direct Submission  
JOURNAL Submitted (20-JAN-2003) Guthrie cDNA Resource Center, Guthrie  
Institute, 1 Guthrie Square, Sayre, PA 18840, USA

FEATURES Location/Qualifiers

1..1068  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1..1068  
/product="chemokine receptor 3"  
1..1068  
/note="g-protein coupled receptor, C-C motif, CCR3"  
/codon\_start=1  
/product="chemokine receptor 3"  
/protein\_id="AA065970.1"  
/db\_xref="GI:29169289"

CDS

mRNA

/translation="MTSLDVTETFGTTSYYDDVGLCEKADTRALMAQFVPLYSIV  
FTVGLGNVVVVMILIKYRLRIWNTNLYLNLASDLLFLVLPFWHYGNHVRG  
HGMCKLSGFHTGLYSIFILITIDRYLAIVHAFALRATVFGVTSITWGL  
AVLAALPEPIFYETEELPEETLCSALYPEDVYSWRHPTLRMTIFGLVPLVMAIC  
VTGIIKTLRCPSKKYKARLIFVIMAVFPIFTVPVVALILSSYQILGNDCCRS  
KHLDLVNLVTEVIAVSHCCANPVYAVGSRFRKYLRFHFRHLLMLHGRYIPLPSR  
KLRTSSVSPTAPELSIVF"

BASE COUNT 231 a 288 c 243 g 306 t  
ORIGIN

Query Match 100.0%; Score:19; DB 9; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 14;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
105 CTGGGCCATCAGTCTCTG 87

RESULT 16  
AR278855/c

LOCUS AR278855 1071 bp mRNA linear PAT 10-APR-2003  
DEFINITION Sequence 6 from patent US 6512103.  
ACCESSION AR278855  
VERSION AR278855.1 GI:29713353

KEYWORDS

SOURCE Unknown.

ORGANISM

Unclassified.

1 (bases 1 to 1071)

Dairaghi, D.J., Hara, T., Miyajima, A., Schall, T.J., Wang, W. and  
Yoshimura, A.

TITLE Mammalian chemokine reagents

JOURNAL Patent: US 6512103-A 6 28-JAN-2003;

FEATURES Location/Qualifiers

1..1071  
/organism="unknown"

BASE COUNT 231 a 292 c 242 g 306 t  
ORIGIN

Query Match 100.0%; Score:19; DB 6; Length 1071;  
Best Local Similarity 100.0%; Pred. No. 14;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
105 CTGGGCCATCAGTCTCTG 87

RESULT 17  
BD015209/c

LOCUS BD015209 1071 bp DNA linear PAT 27-AUG-2002  
DEFINITION Mammalian chemokine CCR18 and receptor CCR3 of mammal.  
ACCESSION BD015209  
VERSION BD015209.1 GI:22556016

KEYWORDS

UNPUBLISHED

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

unclassified

PC C12N15/09, C07K14/715, C07K16/28, C07K19/00, C12P21/02, C12P21/08//  
 PC C12N5/10,  
 PC C12N15/00, C12N5/00  
 CC Strandedness: Single;  
 CC Topology: Linear;  
 CC Mammalian chemokine CCR18 and receptor CCR3 of mammal FH  
 Key Location/Qualifiers  
 FT CDS 1..1071.  
 FEATURES  
 source 1..1071  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 BASE COUNT 231 a 292 c 242 g 306 t  
 ORIGIN  
 Query Match 100.0%; Score 19; DB 6; Length 1071;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCGCATCAGTGTCTG 19  
 Db 105 CTGGGCGCATCAGTGTCTG 87

RESULT 18  
 AR300123/c 1116 bp DNA linear PAT 12-JUN-2003  
 LOCUS AR300123  
 DEFINITION Sequence 5 from patent US 6537764.  
 ACCESSION AR300123  
 VERSION AR300123.1 GI:31687432  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unidentified.  
 REFERENCE 1 (bases 1 to 1116)  
 AUTHORS Gerard, C.J., Gerard, N.P., Mackay, C.R., Ponath, P.D., Post, T.W. and Qin, S.  
 TITLE Method of identifying inhibitors of C-C chemokine receptor 3  
 JOURNAL Patent: US 6537764-A 5 25-MAR-2003;  
 FEATURES Location/Qualifiers  
 source 1..1116  
 /organism="unknown"  
 BASE COUNT 246 a 286 c 257 g 306 t 21 others  
 ORIGIN  
 Query Match 100.0%; Score 19; DB 6; Length 1116;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCGCATCAGTGTCTG 19  
 Db 119 CTGGGCGCATCAGTGTCTG 101

RESULT 19  
 BD082062/c 1116 bp DNA linear PAT 27-AUG-2002  
 LOCUS BD082062  
 DEFINITION G-protein coupled receptor antagonists.  
 ACCESSION BD082062  
 VERSION BD082062.1 GI:22627672  
 KEYWORDS JP 2001524068-A/3.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 1116)  
 REFERENCES Mackay, C.R. and Ponath, P.D.  
 AUTHORS G-protein coupled receptor antagonists  
 TITLE Patent: JP 2001524068-A 3 27-NOV-2001;  
 JOURNAL LEUKOSITE INC  
 COMMENT PN JP 2001524068-A/3

PD 27-NOV-2001  
 PR 24-SEP-1997 JP 1998516642  
 PR 30-SEP-1996 US 08/720565  
 PI CHARLES R MACKAY, PAUL D PONATH  
 PC C07K16/18, C07K16/28, A61K39/395  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 Key Location/Qualifiers.  
 FEATURES  
 source 1..1116  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4577"  
 BASE COUNT 246 a 286 c 257 g 306 t 21 others  
 ORIGIN  
 Query Match 100.0%; Score 19; DB 6; Length 1116;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCGCATCAGTGTCTG 19  
 Db 119 CTGGGCGCATCAGTGTCTG 101

RESULT 20  
 AR300122/c 1193 bp mRNA linear PAT 12-JUN-2003  
 LOCUS AR300122  
 DEFINITION Sequence 3 from patent US 6537764.  
 ACCESSION AR300122  
 VERSION AR300122.1 GI:31687431  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unidentified.  
 REFERENCE 1 (bases 1 to 1193)  
 AUTHORS Gerard, C.J., Gerard, N.P., Mackay, C.R., Ponath, P.D., Post, T.W. and Qin, S.  
 TITLE Method of identifying inhibitors of C-C chemokine receptor 3  
 JOURNAL Patent: US 6537764-A 3 25-MAR-2003;  
 FEATURES Location/Qualifiers  
 source 1..1193  
 /organism="unknown"  
 BASE COUNT 274 a 310 c 275 g 334 t  
 ORIGIN  
 Query Match 100.0%; Score 19; DB 6; Length 1193;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCGCATCAGTGTCTG 19  
 Db 196 CTGGGCGCATCAGTGTCTG 178

RESULT 21  
 BD082061/c 1193 bp DNA linear PAT 27-AUG-2002  
 LOCUS BD082061  
 DEFINITION G-protein coupled receptor antagonists.  
 ACCESSION BD082061  
 VERSION BD082061.1 GI:22627671  
 KEYWORDS JP 2001524068-A/2.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 1193)  
 REFERENCES Mackay, C.R. and Ponath, P.D.  
 AUTHORS G-protein coupled receptor antagonists  
 TITLE Patent: JP 2001524068-A 2 27-NOV-2001;  
 JOURNAL LEUKOSITE INC  
 COMMENT PN JP 2001524068-A/2

	FEATURES	SOURCE
PD	27-NOV-2001	
PF	24-SEP-1997 JP 1998516642	
PR	30-SEP-1996 US 08/720565	
PI	CHARLES R MACKAY, PAUL D PONTTH	
PC	C07K16/18,C07K16/28,A61K39/395	
CC	Strandedness: Double;	
CC	Topology: Linear;	
Key	Location/Qualifiers	
FH	92..115c.	
FT	CDS Location/Qualifiers 1..1193 /organism="Zea mays" /mol_type="genomic DNA" /db_xref="taxon:4577"	
BASE COUNT	274 a     310 c     275 g     334 t	
ORIGIN		

[illegible]

FEATURES		Location/Qualifiers
SOURCE	1. .1201	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
BASE COUNT	278 a 320 c 267 g 336 t	
ORIGIN		
Query Match	100.0%;	Score 19; DB 6; Length 1201;
Best Local Similarity	100.0%;	Pred. No. 14;
Matches 19; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Oy	1 CTGGGCATCAGTGTCTCTG 19	
Db	136 CTGGGCATCAGTGTCTCTG 118	

RESULT	24
LOCUS	HSU28694/c
DEFINITION	Homo sapiens eosinophil CC chemokine receptor 3 mRNA, complete cde.
ACCESSION	U28694
VERSION	U28694.1
KEYWORDS	GI:1199579
SOURCE	.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1200) Combiadere,C., Ahuja,S.K. and Murphy,P.M. Cloning and functional expression of a human eosinophil CC chemokine receptor J. Biol. Chem. 270 (28), 16491-16494 (1995) 95348056 7622448 2 (bases 1 to 1201) Combiadere,C. Direct Submission Submitted (07-JUN-1995) Christophe Combiadere, NIAID, National Institutes of Health, Building 10, Room 1N111, Bethesda, MD 20892, USA 3 (bases 1 to 1201) Combiadere,C., Ahuja,S.K. and Murphy,P.M. Cloning and functional expression of a human eosinophil CC chemokine receptor J. Biol. Chem. 271 (18), 11034 (1996) 96210048 8631926 On Feb 22, 1996 this sequence version replaced gi:861569. [Erratum J. Biol. Chem. 270 (1995) 30235].
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cell_type="eosinophil" 32..1099 /codon_start=1 /product="CC chemokine receptor 3" /protein_id="AAC50469.1" /db_xref="GI:1199580" /translation="MTSTLDVTVEFGTTSYYDDGLLCERADTRALMAQFVPPPLYSLV FTYGLGNNVVVMILIKYRLRIWMNIYLNLAIISDLPLVTLPEFMITHYRGHMVFEG HGKMLISGFPHGISLSEIFPIILLITDRILAIVHAVALRARTVEGVITSIVTWGIC AVLAALDEPFIPTREBELFEETLSAALPEPTVYSMRPHLLPMIIPCLNPLMYATC YTGIITKLRCPSKKTKAIRLRFLVNIAVFIPTPNVAAILISSYSTIIIGNDCCBS KHDLWLVTEVIAYSHSCNMNPVIYAVERFRKYTLHPFHRLMLAGRIPLPSEL KLERTSSVSPTAEPELSIVF"
COMMENT	
JOURNAL	
MEDLINE	
PUBMED	
BASE COUNT	278 a 320 c 267 g 336 t
ORIGIN	
Query Match	100.0%; Score 19; DB 9; Length 1201;
Best Local Similarity	100.0%; Pident. No. 14;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTCTCTG 19  
|||||  
136 CTGGCCATCAGTCTCTG 118

RESULT 25  
AF262300/c  
LOCUS AF262300.1  
DEFINITION Homo sapiens clone 2 CC chemokine receptor 3 (CCR3) mRNA, partial cds.  
ACCESSION AF262300  
VERSION AF262300.1  
KEYWORDS GI:19171642  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1310)  
AUTHORS Vijh,S., Dayhoff,D.E., Wang,C.E., Imam,Z., Ehrenberg,P.K. and Michael,N.L.  
TITLE Transcription Regulation of Human Chemokine Receptor CCR3: Evidence for a Rare TATA-less Promoter Structure Conserved between Drosophila and Humans  
JOURNAL Genomics 80 (1), 86-95 (2002)  
MEDLINE 22074933  
PUBMED 12079287  
REFERENCE 2 (bases 1 to 1310)  
AUTHORS Vijh,S., Dayhoff,D.E., Wang,C.E., Ehrenberg,P.K. and Michael,N.L.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAR-2000) Dept. of Mol. Diag. and Pathogenesis, Walter Reed Army Institute of Research, 1600 E. Gude Drive, Rockville, MD 20850, USA

FEATURES  
source  
1. .1310  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p21"  
/clone="2"  
/cell\_type="Primary eosinophil"  
1. .>1310  
/gene="CCR3"  
1104. .>1310  
/gene="CCR3"  
/note="G-protein coupled receptor; principal eotaxin receptor expressed on eosinophils, CD4 Th2 lymphocytes, CCR lymphocytes, microglia, dendritic cells, and monocytes"  
/codon\_start=1  
/product="CC chemokine receptor 3"  
/protein\_id="AAL85629.1"  
/db\_xref="GI:19171643"  
/translation="MTSLDTVETFGNISYYDVGLCEKADTRALMAQVPLPLSLVFTYGLGNVVMVILIKIRLRIRMT"

BASE COUNT 350 a 317 c 260 g 383 t

ORIGIN

Query Match 100.0%; Score 19; DB 9; Length 1310;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTCTCTG 19  
|||||  
1208 CTGGCCATCAGTCTCTG 1190

Db

RESULT 26  
AR300121/c  
LOCUS AR300121.1  
DEFINITION Sequence 1 from patent US 6537764.

ACCESSION AR300121  
VERSION AR300121.1  
KEYWORDS GI:31687430  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1689)  
AUTHORS Gerard,C.J., Gerard,N.P., Mackay,C.R., Ponath,P.D., Post,T.W. and Qin,S.  
TITLE Method of identifying inhibitors of C-C chemokine receptor 3  
JOURNAL Patent: US 6537764-A 1 25-MAR-2003;  
FEATURES Location/Qualifiers  
source  
1. .1689  
/organism="unknown"  
BASE COUNT 431 a 416 c 344 g 497 t 1 others

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 1689;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTCTCTG 19  
|||||  
285 CTGGCCATCAGTCTCTG 267

Db

RESULT 27  
BD082060/c  
LOCUS BD082060  
DEFINITION G-protein coupled receptor antagonists.  
ACCESSION BD082060  
VERSION BD082060.1  
KEYWORDS GI:22627670  
SOURCE JP 2001524068-A/1.  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 1689)  
AUTHORS Mackay,C.R. and Ponath,P.D.  
TITLE G-protein coupled receptor antagonists  
JOURNAL Patent: JP 2001524068-A 1 27-NOV-2001;  
COMMENT LEUKOSITE INC  
PN JP 2001524068-A/1  
PD 27-NOV-2001  
PF 24-SEP-1997 JP 1998516642  
PR 30-SEP-1996 US 08/720565  
PI CHARLES, R MACKAY, PAUL D PONATH  
PC C07K16/18, C07K16/28, A61K39/395  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
FEATURES  
source  
1. .1689  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"

BASE COUNT 431 a 416 c 344 g 497 t 1 others

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 1689;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTCTCTG 19  
|||||  
285 CTGGCCATCAGTCTCTG 267

Db

RESULT 28  
HSU49727/c  
LOCUS HSU49727  
DEFINITION Human C-C chemokine receptor 3 (CCR-3) gene, complete cds.

ACCESSION U99727  
VERSION U99727.1 GI:1477560  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1689)  
AUTHORS Ponath, P.D., Qin, S., Post, T.W., Wang, J., Wu, L., Gerard, N.P., Newman, W., Gerard, C. and Mackay, C.R.  
TITLE Molecular cloning and characterization of a human eosinophil receptor expressed selectively on eosinophils  
JOURNAL J. Exp. Med. 183 (6), 2437-2448 (1996)  
MEDLINE 96281895  
PUBMED 8676064  
REFERENCE 2 (bases 1 to 1689)  
AUTHORS Ponath, P.D.  
TITLE Direct Submission  
JOURNAL Submitted (21-FEB-1996) Paul D. Ponath, Molecular Biology, Leukosite, Inc., 215 First St., Cambridge, MA 02118, USA  
FEATURES  
source  
1. 1689  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1. 1689  
/gene="CCR-3"  
181. 1248  
/gene="CCR-3"  
/function="G-protein coupled receptor for ectaxin, RANTES and MCP-3"  
/codon\_start=1  
/product="C-C chemokine receptor 3"  
/protein\_id="AAB09726.1"  
/db\_xref="GI:1477561"  
/translation="MTTSLDTVERFTGTSYYDDVGLCEKADTALMAQVPPYLSLVFTVGLGNVVVMMILIKYRLRIMTNLYLNLASDLLFVTLFPWTHYVGNWVGHGKCKLSGFYHGLYSEIFPIILTDRLVIAVHAFALRARTVGVTSIYTWGLAVLAALPEPIFYETEEELFEETLCSALYEDPTVSMRPHRLRMFTICLVPLVMAICVTGIIKTLRCPSKKKYKAIKRLIFVIMAVPFIPTPTPNVAILSSYSIIFGNDCEFTKHLDELVNLVTEVIAVSHCCNMPVIVAVGGRFRKYLRFHRLMLHGRITPLPSBKLERTSSVSSTABPELSIVF"  
BASE COUNT 430 a 416 c 345 g 497 t 1 others  
ORIGIN  
Query Match 100.0%; Score 19; DB 9; Length 1689;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CDS  
1 CTGGGCGCATCAGTCTCTG 19  
|||  
Db 285 CTGGGCGCATCAGTCTCTG 267  
OY  
RESULT 29  
AX334894 1717 bp DNA linear PAT 09-JAN-2002  
LOCUS AX334894 Sequence 5403 from Patent WO0194629.  
DEFINITION AX334894  
ACCESSION AX334894  
VERSION AX334894.1 GI:18125613  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrikan, S., Soppe, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature gene sets  
JOURNAL Patent: WO 0194629-A 5403 13-DEC-2001;  
FEATURES  
location/Qualifiers

source  
1. 1717  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
434 a 428 c 351 g 504 t  
BASE COUNT 434 a 428 c 351 g 504 t  
ORIGIN  
Query Match 100.0%; Score 19; DB 6; Length 1717;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CDS  
1 CTGGGCGCATCAGTCTCTG 19  
|||  
Db 309 CTGGGCGCATCAGTCTCTG 291  
OY  
RESULT 30  
HSU51241/c 1717 bp DNA linear PRI 09-OCT-1996  
LOCUS HSU51241 Human eosinophil ectaxin receptor (CMKBR3) gene, complete cds.  
DEFINITION U51241  
ACCESSION U51241  
VERSION U51241.1 GI:1480480  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1717)  
AUTHORS Daugherty, B.L., Siciliano, S.J., DeMartino, J.A., Malkowitz, L., Strotina, A. and Springer, M.S.  
TITLE Cloning, expression, and characterization of the human eosinophil ectaxin receptor  
JOURNAL J. Exp. Med. 183 (5), 2349-2354 (1996)  
MEDLINE 96235044  
PUBMED 8642344  
REFERENCE 2 (bases 1 to 1717)  
AUTHORS Daugherty, B.L.  
TITLE Direct Submission  
JOURNAL Submitted (13-MAR-1996) Bruce L. Daugherty, Inflammation Research, Merck Research Laboratories, R80W-107, P.O. Box 2000, Rahway, NJ 07065, USA  
FEATURES  
source  
1. 1717  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
205. 1272  
/gene="CMKBR3"  
205. 1272  
/gene="CMKBR3"  
/note="CC chemokine receptor-3; CCR3"  
/codon\_start=1  
/product="eosinophil ectaxin receptor"  
/protein\_id="AAB16831.1"  
/db\_xref="GI:1480481"  
/translation="MTTSLDTVERFTGTSYYDDVGLCEKADTALMAQVPPYLSLVFTVGLGNVVVMMILIKYRLRIMTNLYLNLASDLLFVTLFPWTHYVGNWVGHGKCKLSGFYHGLYSEIFPIILTDRLVIAVHAFALRARTVGVTSIYTWGLAVLAALPEPIFYETEEELFEETLCSALYEDPTVSMRPHRLRMFTICLVPLVMAICVTGIIKTLRCPSKKKYKAIKRLIFVIMAVPFIPTPTPNVAILSSYSIIFGNDCEFTKHLDELVNLVTEVIAVSHCCNMPVIVAVGGRFRKYLRFHRLMLHGRITPLPSBKLERTSSVSSTABPELSIVF"  
BASE COUNT 434 a 428 c 351 g 504 t  
ORIGIN  
Query Match 100.0%; Score 19; DB 9; Length 1717;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CDS  
1 CTGGGCGCATCAGTCTCTG 19  
|||  
Db 309 CTGGGCGCATCAGTCTCTG 291  
OY



```

RESULT 31
LOCUS BD006761/c 1915 bp DNA linear PAT 31-JAN-2002
DEFINITION Chemokine receptors 88-28 [CCR-3] and 88C and antibodies thereof.
ACCESSION BD006761.1 GI:18635132
VERSION JP 2001029089-A/2.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 1915)
AUTHORS Gary, P.W., SheiChart, V.L. and Rayport, C.J.
TITLE Chemokine receptors 88-28 [CCR-3] and 88C and antibodies thereof
JOURNAL Patent: JP 2001029089-A 2 06-FEB-2001;
ICOS CORP
COMMENT OS Unidentified
PN JP 2001029089-A/2
PD 06-FEB-2001
PF 16-MAY-2000 JP 2000143832
PR 20-DEC-1995 US 08/575967, 07-JUN-1996 US 08/661393 PI
PATRICK W GARY, VICKI L SHEICART, CARROLL J RAYPORT PC
C12N15/09, C07K14/715, C07K16/24, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12N5/10, C12N15/02, C12P21/02, C12P21/08, C12Q1/68, G01N33/15, PC
G01N33/50,
PC G01N33/53, G01N33/566//A61K39/395, A61K39/395, A61K45/00, A61P7/02, PC
A61P17/06,
PC A61P19/02, A61P29/00, A61P31/12, (C12P21/02, C12R1.91), C12N15/00,
PC C12N5/00,
PC C12N5/00, C12N15/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1.1915
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 488 a 470 c 373 g 584 t
ORIGIN
Query Match 100.0%; Score 19; DB 6; Length 1915;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTGTCTG 19
Db 466 CTGGCCATCAGTGTCTG 448

RESULT 32
LOCUS BD017703 1915 bp DNA linear PAT 27-AUG-2002
DEFINITION Chemokine receptors 88-28 [CCR-3] and 88C, and antibodies thereof.
ACCESSION BD017703.1 GI:22558879
VERSION JP 2001264324-A/2.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 1915)
AUTHORS Gary, P.W., SheiChart, V.L. and Raport, C.J.
TITLE Chemokine receptors 88-28 [CCR-3] and 88C, and antibodies thereof
JOURNAL Patent: JP 2001264324-A 2 26-SEP-2001;
ICOS CORP
COMMENT OS Unidentified
PN JP 2001264324-A/2
PD 26-SEP-2001
PF 28-DEC-2000 JP 2000401708
PR 20-DEC-1995 US 08/575967, 07-JUN-1996 US 08/661393 PI

```

```

PATRICK W GARY, VICKI L SHEICART, CARROLL J RAYPORT PC
G01N33/50, A61K38/00, A61K45/00, A61P31/12, A61P31/18, C12Q1/02, PC
C12Q1/70,
PC G01N33/15, G01N33/566, G01N33/569//C07K14/705, C07K16/28, C12N15/
PC 09, C12P21/02
PC C12P21/08, (C12Q1/02, C12R1.91), (C12Q1/70, C12R1.93), (C12P21/02,
PC C12R1.91),
PC A61K37/02, C12N15/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1.1915
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 488 a 470 c 373 g 584 t
ORIGIN
Query Match 100.0%; Score 19; DB 6; Length 1915;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTGTCTG 19
Db 466 CTGGCCATCAGTGTCTG 448

RESULT 33
LOCUS BC033514/c 2030 bp mRNA linear PRI 17-OCT-2002
DEFINITION Homo sapiens, similar to chemokine (C-C motif) receptor 3, clone
MGC:34625 IMAGE:5176960, mRNA, complete cds.
ACCESSION BC033514.1 GI:23958837
VERSION BC033514
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2030)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gundaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louised, H.,
Kowle, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAC Plate: 51 Row: n Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4502636.
FEATURES
source
1.2030
/organism="Homo sapiens"
/mol_type="mRNA"

```

```
/db_xref="taxon:9606"
/clone="MGC:34625 IMAGE:5176960"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_id="NIH_MGC_115"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
544.1611
/codon_start=1
/product="similar to chemokine (C-C motif) receptor 3"
/protein_id="AAH3514.1"
/db_xref="GI:23958838"
/db_xref="LocusID:1232"
/translation="MTSLDIVERFGTTSYDDVGLCEKADTRALMAQFVPLYSIV
FTVGLGNVVVMMILIKYRRLRIMTNLYLNLAISDLFLVTLPFWIHVYRGHWVG
HGCKLISGFYHTGLYSIFPIILLTIDRYLAIVHAFALRARTVGTSTVYTWGL
AVLAALPEPIFYETBELFEETLCSALYBEDTVSWRHFHLRMTIFCLVPLVMAIC
YTGIIKTLRCPKSKKTKARLRFVIMAVPFIFWTPYNAVILLSYOSILFGNDCRS
KHLDLWLVTEVIAVSHCCMPVIYAVGGRFRKYLHFFRHLMLHGRYIPPLPSE
KLERTSSVSPSTAPELSIYF"

BASE COUNT      537 a      519 c      446 g      528 t
ORIGIN
Query Match      100.0%; Score 19; DB 9; Length 2030;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 CTGGGCGCATCAGTCTCTG 19
|||||
Db      648 CTGGGCGCATCAGTCTCTG 630

RESULT 34
AX705064/c      5791 bp      DNA      linear      PAT 04-APR-2003
LOCUS      Sequence 5 from Patent WO03014153.
AX705064
VERSION      AX705064.1 GI:29561706
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1
AUTHORS      Renzi,P., Zemzoumi,K. and lamkhioned,B.
TITLE      Cellular virus receptors and methods of use
JOURNAL      Patent: WO 03014153-A 5 20-FEB-2003;
Topigen Pharmaceutical Inc (CA)
FEATURES
source      Location/Qualifiers
1..5791
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
4015..5082
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD88095.1"
/db_xref="GI:29561707"
/translation="MTSLDIVERFGTTSYDDVGLCEKADTRALMAQFVPLYSIV
FTVGLGNVVVMMILIKYRRLRIMTNLYLNLAISDLFLVTLPFWIHVYRGHWVG
HGCKLISGFYHTGLYSIFPIILLTIDRYLAIVHAFALRARTVGTSTVYTWGL
AVLAALPEPIFYETBELFEETLCSALYBEDTVSWRHFHLRMTIFCLVPLVMAIC
YTGIIKTLRCPKSKKTKARLRFVIMAVPFIFWTPYNAVILLSYOSILFGNDCRS
KHLDLWLVTEVIAVSHCCMPVIYAVGGRFRKYLHFFRHLMLHGRYIPPLPSE
KLERTSSVSPSTAPELSIYF"

BASE COUNT      1624 a      1285 c      1160 g      1722 t
ORIGIN
Query Match      100.0%; Score 19; DB 6; Length 5791;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 CTGGGCGCATCAGTCTCTG 19
|||||
```

```
Db      4119 CTGGGCGCATCAGTCTCTG 4101

RESULT 35
AF247361/c      5791 bp      DNA      linear      PRI 26-JUN-2002
LOCUS      Homo sapiens CC chemokine receptor 3 (CCR3) gene, complete cds.
AF247361
DEFINITION      AF247361
VERSION      AF247361.1 GI:19110542
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 5791)
AUTHORS      Vajh,S., Dayhoff,D.E., Wang,C.E., Imam,Z., Ehrenberg,P.K. and
Michael,N.L.
TITLE      Transcription Regulation of Human Chemokine Receptor CCR3: Evidence
for a Rate TATA-less Promoter Structure Conserved between
Drosophila and Humans
JOURNAL      Genomics 80 (1), 86-95 (2002)
MEDLINE      22074933
PubMed      12079287
2 (bases 1 to 5791)
REFERENCE      Vajh,S., Dayhoff,D.E., Wang,C.E., Ehrenberg,P.K. and Michael,N.L.
AUTHORS      Direct Submission
TITLE      Submitted (21-MAR-2000) Dept. of Mol. Diag. and Pathogenesis, U.S.
Military HIV Research Program, 1600 E. Gude Drive, Rockville, MD
20850, USA
FEATURES
source      Location/Qualifiers
1..5791
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p21; between CCR1 and CCR5"
/clone="11A5, frag B"
/cell_type="peripheral blood mononuclear cell"
<1..5562
/genes="CCR3"
<2249..5562
/genes="CCR3"
/product="CC chemokine receptor 3"
<4004..5562
/genes="CCR3"
/product="CC chemokine receptor 3"
4015..5082
/genes="CCR3"
/notes="CCR3; G-protein coupled seven transmembrane
spanning receptor; principle cell-surface receptor for
eosinophil family of chemokines; expressed on eosinophils,
CD4/Th2 and CD8 lymphocytes, monocytes, microglia, and
dendritic cells"
/codon_start=1
/product="CC chemokine receptor 3"
/protein_id="AA85154.1"
/db_xref="GI:19110543"
/translation="MTSLDIVERFGTTSYDDVGLCEKADTRALMAQFVPLYSIV
FTVGLGNVVVMMILIKYRRLRIMTNLYLNLAISDLFLVTLPFWIHVYRGHWVG
HGCKLISGFYHTGLYSIFPIILLTIDRYLAIVHAFALRARTVGTSTVYTWGL
AVLAALPEPIFYETBELFEETLCSALYBEDTVSWRHFHLRMTIFCLVPLVMAIC
YTGIIKTLRCPKSKKTKARLRFVIMAVPFIFWTPYNAVILLSYOSILFGNDCRS
KHLDLWLVTEVIAVSHCCMPVIYAVGGRFRKYLHFFRHLMLHGRYIPPLPSE
KLERTSSVSPSTAPELSIYF"

BASE COUNT      1624 a      1285 c      1160 g      1722 t
ORIGIN
Query Match      100.0%; Score 19; DB 9; Length 5791;
Best Local Similarity 100.0%; Pred. No. 15;
```

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
Db 4119 CTGGGCCATCAGTCTCTG 4101

RESULT 36  
AF224496S2/c 7010 bp DNA linear PRI 02-MAY-2001

LOCUS AF224496S2 7010 bp DNA linear PRI 02-MAY-2001  
DEFINITION Homo sapiens CC chemokine receptor 3 (CCR3) gene, exon 2 and  
partial cds.  
ACCESSION AF224497  
VERSION AF224497.1 GI:13924486  
KEYWORDS  
SEGMENT 2 of 2  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 7010)  
Scoret, E.J.  
CCR3 expression is associated with chromatin remodeling in Th2  
cells  
AUTHORS  
TITLE Unpublished  
REFERENCE 2 (bases 1 to 7010)  
Scoret, E.J.  
Direct Submission  
JOURNAL Submitted (13-JAN-2000) Basel Institute for Immunology, 487  
Grenzacherstrasse, Basel CH-4005, Switzerland  
FEATURES  
source 1..7010  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
order(AF224496.1:298..957.1..>7010)  
/gene="CCR3"  
join(AF224496.1:298..376.6784..>7010)  
/product="CCR3"  
/product="CC chemokine receptor 3"  
repeat\_region 4318..4570  
/rpt\_family="Alu"  
/rpt\_type="dispersed"  
6784..7010  
/gene="CCR3"  
/number=2  
6795..>7010  
/gene="CCR3"  
/codon\_start=1  
/product="CC chemokine receptor 3"  
/protein\_id="AAK49028.1"  
/db\_xref="GI:13924487"  
/translation="MTTSLDTVEFTGTYSDVGLCEKADTRALMAQFVPPPLYSIV  
FTVGLGNVVVVWILLIKYRRLRIMNTNY"  
BASE COUNT 2136 a 1420 c 1320 g 2129 t 5 others

ORIGIN  
Query Match 100.0%; Score 19; DB 9; Length 7010;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
Db 6899 CTGGGCCATCAGTCTCTG 6881

RESULT 37  
AC138069/c 177334 bp DNA linear PRI 19-FEB-2003

LOCUS AC138069 177334 bp DNA linear PRI 19-FEB-2003  
DEFINITION Homo sapiens chromosome 3 clone RP13-54612, complete sequence.  
ACCESSION AC138069  
VERSION AC138069.3 GI:28416170  
KEYWORDS HTG.

SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 177334)  
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,  
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and  
Haugen, E.D.  
Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 177334)  
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
Direct Submission  
JOURNAL Submitted (12-DEC-2002) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
REFERENCE 3 (bases 1 to 177334)  
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
Direct Submission  
JOURNAL Submitted (10-JAN-2003) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
REFERENCE 4 (bases 1 to 177334)  
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,  
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and  
Haugen, E.D.  
Direct Submission  
JOURNAL Submitted (19-FEB-2003) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Feb 19, 2003 this sequence version replaced gi:27573398.

COMMENT  
Genome Center  
Center: University of Washington Genome Center  
Center Code: UMG  
Web site: <http://www.genome.washington.edu>  
Contact: [uwgchgs@u.washington.edu](mailto:uwgchgs@u.washington.edu)  
----- Project Information  
Center project name: chr-3  
Center clone name: RP13-54612 (bc08020)  
----- Summary Statistics  
Sequencing vector: plasmid; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 177210 bases at least Q40  
Consensus quality: 177313 bases at least Q30  
Consensus quality: 177334 bases at least Q20  
Insert size: 177334; sum-of-contigs  
Quality coverage: 9.3x in Q20 bases; sum-of-contigs

Overlapping Sequences:  
5': RP11-793E15 (UMGC:bc0564) AC104439, 95469-bp overlap  
3': U95626, 42710-bp overlap

Sequence Quality Assessment:  
This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
Genbank flat file format but are available as part  
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:  
all regions were either double-stranded or sequenced with an  
alternate chemistry or covered by high quality data (i.e., Phred  
quality >= 30); an attempt was made to resolve all sequencing  
problems, such as compressions and repeats; all regions were  
covered by at least one plasmid subclone or more than one M13  
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:  
This sequence has been validated by Multiple Complete Digest  
fingerprinting. Comparison of the experimentally derived digest  
fragments with sequence-predicted fragments is given below.  
The electronically-digested sequence consists of both insert and

vector, in order to accurately represent the entire circular BAC.  
Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

## HindIII

## EcoRI

## BglIII

SeqDerMap	FingerPrint	SeqDerMap	FingerPrint	SeqDerMap	FingerPrint
1239	1199	11125	10952	6306	6319
449	<800	2290	2310	2067	2071
510	<800	560	<800	3913	3756
6511	6363	54	<800	2169	2210
5296	5509	1159	1158	910	910
10424	10027	4052	4034	875	883
5046	5119	1846	1829	3716	3756
1196	1199	8334	8273	4502	4523
1397	1392	12882	12573	1967	1945
2597	2624	448	<800	2864	2871
1688	1674	12737	12573	4724	4738
3800	3818	10300	10103	4773	4738
2248	2269	6671	6656	3707	3756
1074	1055	4052	4034	336	<800
7303	7718	1392	1386	79	<800
959	959	4253	4286	4736	4738
551	<800	2791	2834	1889	1945
1074	1055	2647	2682	3696	3756
5688	5759	3423	3409	1569	1539
5183	5363	73	<800	8078	8038
3100	3093	499	<800	1365	1325
921	959	723	<800	289	<800
4350	4244	2763	2834	11047	10790
4931	4857	6594	6656	8848	8963
3907	3818	4503	4468	1914	1945
8001	7718	5154	5136	1336	1325
3804	3818	4730	4713	11485	11341
2355	2376	13628	13324	974	975
862	868	166	<800	2448	2501
1711	1674	9554	9494	373	<800
5672	5759	1398	1386	1003	975

3900	3818	1613	1641	1787	1756
1905	1854	1479	1503	857	883
7872	7718	1540	1503	1974	1945
1174	1155	520	<800	1707	1631
7015	7139	219	<800	3835	3756
2852	2868	1571	1503	6586	6551
221	<800	154	<800	178	<800
1002	1055	5077	5136	2905	2871
201	<800	2156	2190	2500	2501
6717	6864	1075	1089	389	<800
237	<800	162	<800	41	<800
2737	2749	2752	2834	685	<800
3448	3416	1647	1641	2598	2662
3256	3238	7059	7056	248	<800
5801	5759	1547	1503	2152	2210
1512	1497	5163	5136	431	<800
1398	1392	8679	8719	237	<800
8413	8333	2033	2035	3380	3411
6621	6614	886	892	5694	5744
645	<800			3255	3265
1026	1055			183	<800
14333	14484			3128	3144
				10701	10790
				1321	1325

Query Match 100.0%; Score 19; DB 9; Length 177334;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCGCATCAGTCTCTG 19  
DB 88495 CTGGGCGCATCAGTCTCTG 88477

RESULT 38  
AC104439/c 197279 bp DNA linear PRI 20-JUN-2002  
LOCUS AC104439  
DEFINITION Homo sapiens chromosome 3 clone RP11-793E15, complete sequence.  
AC104439 AC024739  
VERSION AC104439.2 GI:21490240  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 197279)  
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,  
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.

TITLE and Haugen, E.D.  
JOURNAL Direct Submission  
REFERENCE Unpublished  
AUTHORS 2 (bases 1 to 197279)  
TITLE Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
JOURNAL Direct Submission  
Submitted (11-DEC-2001) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
3 (bases 1 to 197279)  
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,  
Szenhimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.  
and Haugen, E.D.  
JOURNAL Direct Submission  
Submitted (20-JUN-2002) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Jun 20, 2002 this sequence version replaced gi:17488621.  
COMMENT ----- Genome Center  
Center: University of Washington Genome Center  
Center Code: UWGC  
Web site: <http://www.genome.washington.edu>  
Contact: uwgchgs@u.washington.edu  
Drafting Center: WUGSC  
----- Project Information  
Center project name: Chr-3  
Center clone name: Rp11-793B15 (bc0564)  
----- Summary Statistics  
Sequencing vector: unknown; 52% of reads  
Chemistry: Dye-terminator ET; 94% of reads  
Chemistry: Dye-terminator Big Dye; 6% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 197168 bases at least Q40  
Consensus quality: 197255 bases at least Q30  
Consensus quality: 197275 bases at least Q20  
Insert size: 197279; sum-of-contigs  
Quality coverage: 8.2x in Q20 bases; sum-of-contigs  
-----  
Overlapping Sequences:  
5' : Rp11-91E8 (UWGC:bc0216) AC026349  
3' : CTD-2563A18 (UWGC:bc0730)  
-----  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.  
-----  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.  
-----  
Sequence Validation:  
This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.  
HindIII BglII EcoRI  
SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt

2687	2617	8949	8586	8696	8661
6382	6410	2067	2160	6	<800
512	<800	7846	7940	2742	2803
449	<800	3734	3895	5376	5324
2602	2763	1334	1301	1493	1478
2590	2617	2287	2309	823	835
8313	8291	1814	1918	1962	2002
1711	1683	691	<800	2900	2954
9821	9472	5477	5348	1484	1478
516	<800	305	<800	1005	995
8587	8291	25245	25541	1181	1171
7446	7581	3988	4121	18560	19002
2088	2075	1633	1598	3603	3579
2509	2617	631	<800	4943	5076
3519	3501	90	<800	3239	3241
26	<800	402	<800	953	995
925	933	3350	3490	1621	1615
98	<800	4577	4515	6827	6900
6409	6410	1229	1183	79	<800
1383	1376	4221	4331	16418	16263
16048	15829	3980	4121	3361	3241
1876	1877	2548	2497	872	881
4169	4068	674	<800	4220	4193
1674	1683	2332	2309	2848	2803
52	<800	11445	11045	2283	2318
5227	5082	5692	5632	3211	3241
15464	15829	1385	1414	4615	4632
14333	14286	608	<800	6823	6900
1026	1026	3930	3895	886	881
645	<800	1465	1414	2032	2002
6621	6410	2382	2497	8660	8661
8433	8291	747	<800	5163	5076
1401	1376	1951	2012	1547	1478
1512	1471	642	<800	7058	6900
5801	5644	21060	21003	1647	1615

3256	3278	3700	3666	2753	2803
3448	3501	1331	1301	162	<800
2738	2763	10705	10502	1078	1093
237	<800	3128	3317	2156	2173
6715	6823	183	<800	5077	5076
201	<800	3255	3490	154	<800
1002	1026	5633	5632	1571	1478
221	<800	3380	3317	219	<800
2854	2899	237	<800	520	<800
6998	7317	431	<800	1540	1478
1174	1145	2152	2160	1479	1478
7876	8291	249	<800	1613	1615
1911	1877	2599	2778	1399	1478
3899	3847	685	<800	9553	9383
5672	5644	41	<800	166	<800
1711	1683	388	<800	13615	13115
862	854	2500	2497	4731	4632
2368	2389	2889	2778	5159	5076
		178	<800	4503	4430
		6589	6728	6607	6900
		3836	3895	2766	2803

Query Match 100.0%; Score 19; DB 9; Length 197279;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
 DB 190292 CTGGGCCATCAGTCTCTG 190274

RESULT 39  
 HSA12688/c 220965 bp DNA linear HTG 15-MAY-2002  
 LOCUS HSA12688 Homo sapiens chromosome 3 clone RP6-32g23 map 3p21.3, \*\*\*  
 DEFINITION SEQUENCING IN PROGRESS \*\*\*, 26 ordered pieces.  
 ACCESSION AJ12688.2 GI:13559235  
 VERSION AJ12688  
 KEYWORDS HTG; HTGS; PHASE2.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 Kiss, H., Yang, Y., Kiss, C., Andersson, K., Klein, G., Imreh, S. and  
 Dunanski, J.P.  
 TITLE The transcriptional map of the common eliminated region 1 (C3CR1)  
 in 3p21.3  
 JOURNAL Eur. J. Hum. Genet. 10 (1), 52-61 (2002)  
 MEDLINE 21906202  
 PUBMED 11896456  
 REFERENCE 2 (bases 1 to 220965)

# AUTHORS TITLE JOURNAL COMMENT

Kiss, H.  
 Direct Submission  
 Submitted (01-APR-2001) Kiss H., Microbiology and Tumorbiology  
 Center (MTC), Karolinska Institute, Box 280, Stockholm, S-17177,  
 SWEDEN  
 On Apr 5, 2001 this sequence version replaced gi:13548633.  
 The sequence is a consensus sequence of clone RP4-787c23 (1-140400  
 bp),  
 clone RP6-32g23 (31212-220965 bp), clone RP6-146e1 (partially,  
 1-6800 bp)  
 and clone RP6-188g11 (partially, 1-108303 bp). The sequencing  
 contigs are in order and the gaps between them are represented by 100 Ns.  
 Contig 1:  
 1-11731 bp Contig 2: 11832-26218 bp Contig 3: 26319-28347 bp  
 Contig 4:  
 28448-42160 bp Contig 5: 42261-55059 bp Contig 6: 55160-61578 bp  
 Contig  
 7: 61679-97342 bp Contig 8: 97443-117655 bp Contig 9:  
 117756-118727 bp  
 Contig 10: 118828-121834 bp Contig 11: 121935-127855 bp Contig  
 12:  
 127856-129383 bp Contig 13: 129484-131747 bp Contig 14:  
 131848-132316 bp  
 Contig 15: 132417-134455 bp Contig 16: 134556-135527 bp Contig  
 17:  
 135628-189051 bp Contig 18: 189152-189476 bp Contig 19:  
 189577-191375 bp  
 Contig 20: 191476-201473 bp Contig 21: 201574-202307 bp Contig  
 22:  
 202408-204878 bp Contig 23: 204979-213531 bp Contig 24:  
 213632-218109 bp  
 Contig 25: 218210-219800 bp Contig 26: 219901-220965 bp.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 26 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 11731: contig of 11731 bp in length  
 \* 11732 11831: gap of 100 bp  
 \* 11832 26218: contig of 14387 bp in length  
 \* 26219 26318: gap of 100 bp  
 \* 26319 28347: contig of 2029 bp in length  
 \* 28348 28448: gap of 100 bp  
 \* 28448 42160: contig of 13713 bp in length  
 \* 42161 42361: gap of 200 bp  
 \* 42361 55059: contig of 12699 bp in length  
 \* 55060 55159: gap of 100 bp  
 \* 55160 61578: contig of 6419 bp in length  
 \* 61579 61678: gap of 100 bp  
 \* 61678 97342: contig of 35664 bp in length  
 \* 97343 97442: gap of 100 bp  
 \* 97442 117655: contig of 20213 bp in length  
 \* 97443 117755: gap of 100 bp  
 \* 117656 118727: contig of 972 bp in length  
 \* 118728 118827: gap of 100 bp  
 \* 118828 121834: contig of 3007 bp in length  
 \* 121835 121934: gap of 100 bp  
 \* 121935 127855: contig of 5921 bp in length  
 \* 127856 127955: gap of 100 bp  
 \* 127956 129383: contig of 1428 bp in length  
 \* 129384 129483: gap of 100 bp  
 \* 129484 131747: contig of 2264 bp in length  
 \* 131748 131847: gap of 100 bp  
 \* 131848 132316: contig of 469 bp in length  
 \* 132317 132417: gap of 100 bp  
 \* 132417 134455: contig of 2039 bp in length  
 \* 134456 134555: gap of 100 bp  
 \* 134556 135527: contig of 972 bp in length

```

* 135528 135627: gap of 100 bp
* 135628 189051: contig of 53424 bp in length
* 189052 189151: gap of 100 bp
* 189152 189476: contig of 325 bp in length
* 189477 189576: gap of 100 bp
* 189577 191375: contig of 1799 bp in length
* 191376 191475: gap of 100 bp
* 191476 201473: contig of 9998 bp in length
* 201474 201573: gap of 100 bp
* 201574 202307: contig of 734 bp in length
* 202308 202407: gap of 100 bp
* 202408 204878: contig of 2471 bp in length
* 204879 204978: gap of 100 bp
* 204979 213531: contig of 8553 bp in length
* 213532 213631: gap of 100 bp
* 213632 218109: contig of 4478 bp in length
* 218110 218209: gap of 100 bp
* 218210 219800: contig of 1591 bp in length
* 219801 219901: gap of 100 bp
* 219901 220965: contig of 1065 bp in length.

FEATURES
    source
        1..220965
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="3"
            /map="13p21.3"
            /clone="RP6-32g23"

BASE COUNT  62577 a 44674 c 45661 g 65453 t 2600 others
ORIGIN
Query Match      100.0%; Score 19; DB 2; Length 220965;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCATGCTCTG 19
    |||||
Db 140368 CTGGGCCATCATGCTCTG 140350

RESULT 40
BC030031/c 1508 bp mRNA linear ROD 20-SEP-2002
LOCUS Mus musculus, lymphocyte antigen 108, clone MGC:41003
DEFINITION IMAGE:1225313, mRNA, complete cds.
ACCESSION BC030031 GI:20988098
VERSION BC030031.1
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1508)
Strausberg, R.
Direct Submission
Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Marcello Bento Soares, Ph.D.
cDNA Library Preparation: Soares Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whitting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

```

```

Series: IRAC Plate: 66 Row: d Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13507653.

FEATURES
    source
        1..1508
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="MGC:41003 IMAGE:1225313"
            /tissue_type="thymus gland, mouse"
            /clone_lib="Soares_thymus_2NDMT"
            /lab_host="DH10B"
            /note="Vector: pTT73D-Pac"
            153..1208
            /codon_start=1
            /product="lymphocyte antigen 108"
            /protein_id="AAH30031.1"
            /db_xref="GI:20988098"
            /db_xref="LOCUSID:30925"
            /translacion="MAVSRAPAPDSACORWWLFLPVFCIGSGSEVSSSDPOLNNG
VIGESAVIPKILPAKIANIITIMYEMASOVTALVININSPSPQIMNTDVKRLNI
TGSYSIQISNLTMADTGSYTAQITTDSEVITFKTILRPERLGNLETTITLLENG
TQIHILACVLRKQSVSVQWQATGNI SLGPNVTIFMDPRNSGDQTYCRANAVSN
LSVSVSTQSLCKGVLTPNPNNAVFMWTTIISAVLILFVCSIHMKRGSLLPISO
HPESOSTDGPSPGNTVYAQVTRPMQEKIKPKIPINDSMITISIVNHSREETVALTG
VNPITLKNTLILNYS"

BASE COUNT  482 a 342 c 321 g 363 t
ORIGIN
Query Match      91.6%; Score 17.4; DB 10; Length 1508;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCCATCATGCTCTG 19
    |||||
Db 1014 CTGGGCCATCATGCTCTG 996

Search completed: January 13, 2004, 17:57:07
Job time : 747 secs

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 13, 2004, 15:53:18 ; Search time 1192.5 Seconds  
(without alignments)  
387.241 Million cell updates/sec

Title: US-09-719-737-18

Perfect score: 19  
Sequence: 1 ctggggccatcagtcctctg 19

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:  
1: em\_estdb:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estov:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estro:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rnd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vit1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	19	100.0	745 10	BG204024 RST23417
C 2	19	100.0	789 10	BG205056 RST24475
C 3	19	100.0	835 12	B1821796 B1906283
C 4	19	100.0	876 12	B1906283 603063222

Result No.	Score	Query Match Length	ID	Description
C 5	19	100.0	884 10	BG182330 RST1196 A
C 6	19	100.0	934 10	BG460984 RST43688
C 7	17.4	91.6	451 28	AO571744 HS 5380 B
C 8	17.4	91.6	692 14	B758727 B758727
C 9	17.4	91.6	703 12	BB089480 BB089480
C 10	17.4	91.6	737 12	BG821437 BG821437
C 11	17.4	91.6	738 12	BG824723 BG824723
C 12	17.4	91.6	769 12	BG821578 BG821578
C 13	17.4	91.6	805 12	B1836792 B1836792
C 14	17.4	91.6	808 12	BG822063 BG822063
C 15	17.4	91.6	900 12	BG821494 BG821494
C 16	17.4	91.6	908 29	CNS03A4E CNS03A4E
C 17	17.4	91.6	1070 29	CNS03SN1 CNS03SN1
C 18	17.4	91.6	1579 10	BG327989 BG327989
C 19	17.4	91.6	1817 11	AK034665 AK034665
C 20	16.4	86.3	137 6	AL810567 TRL1CUM
C 21	16.4	86.3	178 10	BF956881 BF956881
C 22	16.4	86.3	292 14	CA912316 CA912316
C 23	16.4	86.3	322 10	BF328982 BF328982
C 24	16.4	86.3	327 9	AW594895 AW594895
C 25	16.4	86.3	381 29	CC469335 CC469335
C 26	16.4	86.3	386 14	CB806515 CB806515
C 27	16.4	86.3	422 13	B1596515 B1596515
C 28	16.4	86.3	431 28	AZ786165 AZ786165
C 29	16.4	86.3	460 28	AQ706454 AQ706454
C 30	16.4	86.3	474 28	AZ813275 AZ813275
C 31	16.4	86.3	477 12	B1473885 B1473885
C 32	16.4	86.3	522 10	BG728994 BG728994
C 33	16.4	86.3	528 12	BG891847 BG891847
C 34	16.4	86.3	536 13	BQ616929 BQ616929
C 35	16.4	86.3	574 13	BQ074936 BQ074936
C 36	16.4	86.3	591 12	B1670955 B1670955
C 37	16.4	86.3	596 14	CA500347 CA500347
C 38	16.4	86.3	602 9	A1496684 A1496684
C 39	16.4	86.3	627 13	BQ074707 BQ074707
C 40	16.4	86.3	672 10	BE557791 BE557791
C 41	16.4	86.3	685 12	BM141543 BM141543
C 42	16.4	86.3	685 28	AO995355 AO995355
C 43	16.4	86.3	689 9	AP160693 AP160693
C 44	16.4	86.3	691 9	AW342617 AW342617
C 45	16.4	86.3	700 14	CD237190 CD237190

## ALIGNMENTS

RESULT 1  
BG204024/c 745 bp mRNA linear EST 21-APR-2001  
RST23417 Atherys RAGE library Homo sapiens cDNA, mRNA sequence.  
DEFINITION  
BG204024  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
EST.  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 745)  
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J.,  
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Maye,R., Smith  
,E., Veloso,N., Kilka,A., Hess,J., Cothren,K., Lo,K., Offenbacher  
,J., Danzig,J. and Ducar,M.  
Creation of genome-wide protein expression libraries using random  
activation of gene expression  
Nat. Biotechnol. 19 (5), 440-445 (2001)  
JOURNAL  
MEDLINE  
PUBMED  
11329013  
21227151  
Contact: Scott J. Cain  
Atherys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596

Email: scai@atherys.com  
High quality sequence stop: 447.  
Location/Qualifiers

FEATURES  
Source

1.745  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT 159 a 207 c 170 g 209 t  
ORIGIN

Query Match 100.0%; Score 19; DB 10; Length 745;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGGCCATCAGTCTCTG 19  
|||||  
Db 111 CTGGGCCATCAGTCTCTG 93

RESULT 2  
BG205056 789 bp mRNA linear EST 21-APR-2001  
LOCUS RST24475 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
DEFINITION BG205056  
ACCESSION BG205056  
VERSION BG205056.1 GI:13726743  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 789)  
AUTHORS Harrington,J.O., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J.,  
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,  
E., Veloso,N., Klika,A., Hess,J., Cochran,K., Lo,K., Offenbacher,  
J., Danzig,J. and Ducar,M.

TITLE Creation of genome-wide protein expression libraries using random  
activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave. Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scai@atherys.com

FEATURES  
Source  
High quality sequence stop: 521.  
Location/Qualifiers

1.789  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT 174 a 207 c 183 g 222 t  
ORIGIN

Query Match 100.0%; Score 19; DB 10; Length 789;

Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGGCCATCAGTCTCTG 19  
|||||  
Db 120 CTGGGCCATCAGTCTCTG 102

RESULT 3  
B1821796/c 835 bp mRNA linear EST 04-OCT-2001  
LOCUS B1821796  
DEFINITION 60303587221 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5176960 5',  
RNA sequence.  
ACCESSION B1821796  
VERSION B1821796.1 GI:15933346  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 835)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: L1AM11440 row: P column: 17

High quality sequence start: 2  
High quality sequence stop: 750.  
Location/Qualifiers

FEATURES  
Source  
1.835  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5176960"  
/lab\_host="DH10B"  
/clone\_lib="NIH MGC 115"

/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1  
male lung, age 27; and 1 male testis, age 65. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH\_MGC Library."

BASE COUNT 206 a 217 c 210 g 202 t  
ORIGIN

Query Match 100.0%; Score 19; DB 12; Length 835;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGGCCATCAGTCTCTG 19  
|||||  
Db 650 CTGGGCCATCAGTCTCTG 632

RESULT 4  
B1906283/c 876 bp mRNA linear EST 16-OCT-2001  
LOCUS B1906283  
DEFINITION 60306322221 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5212541 5',  
mRNA sequence.  
ACCESSION B1906283  
VERSION B1906283.1 GI:16168946

Query Match 100.0%; Score 19; DB 12; Length 876;

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>  
Plate: LLM11533 row: K column: 06  
High quality sequence stop: 800.  
Location/Qualifiers

FEATURES  
Source  
1..876  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5212541"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_1lb="NIH MGC 118"  
/note="Vector: PCMV-SPORT6; Site 1: NotI, Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."

BASE COUNT  
202 a 232 c 193 g 249 t

ORIGIN

Query Match 100.0%; Score 19; DB 12; Length 876;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CTGGGCGCATCAGTCTCTG 19  
|||||  
Db 183 CTGGGCGCATCAGTCTCTG 165

RESULT 5  
BG182330 884 bp mRNA linear EST 21-APR-2001  
LOCUS BG182330  
DEFINITION RST1166 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG182330  
VERSION BG182330.1 GI:13704017  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 884)  
AUTHORS Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R., Cain, S., Leventhal, C., Thornton, M., Ramchandran, R., Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith, E., Veloso, N., Kilka, A., Hess, J., Cochren, K., Lo, K., Offenbacher, J., Danzig, J., and Ducar, M.  
TITLE Creation of genome-wide protein expression libraries using random activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain

Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: [scain@athersys.com](mailto:scain@athersys.com)  
High quality sequence stop: 529.  
Location/Qualifiers

FEATURES  
Source  
1..884  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_1lb="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT  
191 a 231 c 206 g 256 t

ORIGIN

Query Match 100.0%; Score 19; DB 10; Length 884;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CTGGGCGCATCAGTCTCTG 19  
|||||  
Db 121 CTGGGCGCATCAGTCTCTG 103

RESULT 6  
BG460984 934 bp mRNA linear EST 21-APR-2001  
LOCUS BG460984  
DEFINITION RST43688 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG460984  
VERSION BG460984.1 GI:13749490  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 934)  
AUTHORS Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R., Cain, S., Leventhal, C., Thornton, M., Ramchandran, R., Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith, E., Veloso, N., Kilka, A., Hess, J., Cochren, K., Lo, K., Offenbacher, J., Danzig, J., and Ducar, M.  
TITLE Creation of genome-wide protein expression libraries using random activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: [scain@athersys.com](mailto:scain@athersys.com).

FEATURES  
Source  
1..934  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_1lb="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT  
200 a 188 c 237 g 282 t





Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM1728 row: k column: 23  
High quality sequence stop: 734.  
Location/Qualifiers

## FEATURES

source

1. 737  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4864654"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 15"  
/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"  
BASE COUNT 170 a 214 c 223 g 130 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 737;  
Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGGCATCAGTCTCTG 19  
|||||  
45 CTGGGCATCAGTCTCTG 27

Db

RESULT 11  
BG824723 738 bp mRNA linear EST 22-MAY-2001  
LOCUS 602728705F1 NIH\_MGC\_15 Homo sapiens CDNA clone IMAGE:4868064 5',  
DEFINITION mRNA sequence.  
ACCESSION BG824723  
VERSION BG824723.1 GI:14172310  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 738)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM1737 row: j column: 01  
High quality sequence stop: 728.  
Location/Qualifiers

## FEATURES

source

1. 738  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4868064"  
/tissue\_type="adenocarcinoma cell line"

/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 15"  
/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"  
BASE COUNT 168 a 211 c 229 g 130 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 736;  
Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGGCATCAGTCTCTG 19  
|||||  
45 CTGGGCATCAGTCTCTG 27

Db

RESULT 12  
BG821578 769 bp mRNA linear EST 22-MAY-2001  
LOCUS 602725172F1 NIH\_MGC\_15 Homo sapiens CDNA clone IMAGE:4864768 5',  
DEFINITION mRNA sequence.  
ACCESSION BG821578  
VERSION BG821578.1 GI:14169165  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 769)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM1728 row: p column: 17  
High quality sequence stop: 692.  
Location/Qualifiers

## FEATURES

source

1. 769  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4864768"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 15"  
/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"  
BASE COUNT 182 a 217 c 235 g 135 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 769;  
Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
Db 45 CTGGGCCATCAGTCTCTG 27

RESULT 13  
LOCUS BI836792 805 bp mRNA linear EST 04-OCT-2001  
DEFINITION 603084870F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5224073 5', mRNA sequence.

ACCESSION BI836792  
VERSION BI836792.1 GI:15948342  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L16M11563 row: k column: 18  
High quality sequence start: 2  
High quality sequence stop: 635.  
Location/Qualifiers

FEATURES  
source  
1..805  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5224073"  
/lab\_host="DH10B"  
/clone\_1lb="NIH\_MGC\_120"  
/note="Organ: pooled pancreas and spleen; Vector: pcwv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

BASE COUNT  
190 a 219 c 245 g 151 t

Query Match 91.6%; Score 17.4; DB 12; Length 805;  
Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
Db 527 CTGGGCCATCAGTCTCTG 545

RESULT 14  
LOCUS BG822063 898 bp mRNA linear EST 22-MAY-2001  
DEFINITION 602726178F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4865803 5', mRNA sequence.

ACCESSION BG822063  
VERSION BG822063.1 GI:14169650  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS 1 (bases 1 to 898)  
TITLE NIH-MGC http://mgs.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L16M1731 row: k column: 20  
High quality sequence stop: 821.  
Location/Qualifiers

FEATURES  
source  
1..898  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4865803"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NIH\_MGC\_15"  
/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGCGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT  
200 a 258 c 278 g 162 t

Query Match 91.6%; Score 17.4; DB 12; Length 898;  
Best Local Similarity 94.7%; Pred. No. 1.2e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
Db 45 CTGGGCCATCAGTCTCTG 27

RESULT 15  
LOCUS BG821494 900 bp mRNA linear EST 22-MAY-2001  
DEFINITION 602725068F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4864577 5', mRNA sequence.

ACCESSION BG821494  
VERSION BG821494.1 GI:14169081  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L16M1728 row: h column: 18  
High quality sequence stop: 808.  
Location/Qualifiers

source

1..900

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_image="4864577"

/issue\_type="adenocarcinoma cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 15"

/note="Organ: colon; Vector: pOTV1; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 211 a 258 c 273 g 158 t

ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 900;

Best Local Similarity 94.7%; Pred. No. 1.2e+03;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCGCATGAGTCTCTG 19

45 CTGGGCGCATGAGTCTCTG 27

Db

RESULT 16

CNS03J4E

LOCUS 908 bp DNA linear GSS 01-SEP-2000

DEFINITION Tetradon nigroviridis genome survey sequence PUC-Orl end of clone 008L02 of library G from Tetradon nigroviridis, genomic survey sequence.

ACCESSION AL234743

VERSION AL234743.1 GI:7893878

KEYWORDS GSS; genome survey sequence.

SOURCE Tetradon nigroviridis

ORGANISM Tetradon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetradon.

REFERENCE 1

AUTHORS Roest Crolius,H., Jallou,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Wincker,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.

TITLE Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence

JOURNAL Nat. Genet. 25 (2), 235-238 (2000)

MEDLINE 20296633

PUBMED 10835645

TITLE 2

AUTHORS Roest Crolius,H., Jallou,O., Dasilva,C., Ozouf-Costaz,C., Fitzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis

JOURNAL Genome Res. 10 (7), 939-949 (2000)

MEDLINE 20359837

PUBMED 10899143

TITLE 3 (bases 1 to 908)

REFERENCE Genoscope.

AUTHORS Direct Submission

TITLE Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT - Web : www.genoscope.cns.fr)

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetradon>.

FEATURES

Source 1..908

Location/Qualifiers

/organism="Tetradon nigroviridis"

/mol\_type="genomic DNA"

/db\_xref="taxon:99883"

/clone="008L02"

/clone\_lib="G"

/note="Genoscope sequence ID : COBG008DF01SP1-end : PUC-Orl"

BASE COUNT 229 a 255 c 195 g 229 t

ORIGIN

Query Match 91.6%; Score 17.4; DB 29; Length 908;

Best Local Similarity 94.7%; Pred. No. 1.2e+03;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCGCATGAGTCTCTG 19

877 CTGGGCGCATGAGTCTCTG 895

Db

RESULT 17

CNS03SN1

LOCUS 1070 bp DNA linear GSS 01-SEP-2000

DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone 054A19 of library G from Tetradon nigroviridis, genomic survey sequence.

ACCESSION AL258742

VERSION AL258742.1 GI:7979754

KEYWORDS GSS; genome survey sequence.

SOURCE Tetradon nigroviridis

ORGANISM Tetradon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetradon.

REFERENCE 1

AUTHORS Roest Crolius,H., Jallou,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Wincker,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.

TITLE Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence

JOURNAL Nat. Genet. 25 (2), 235-238 (2000)

MEDLINE 20296633

PUBMED 10835645

TITLE 2

AUTHORS Roest Crolius,H., Jallou,O., Dasilva,C., Ozouf-Costaz,C., Fitzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis

JOURNAL Genome Res. 10 (7), 939-949 (2000)

MEDLINE 20359837

PUBMED 10899143

TITLE 3 (bases 1 to 1070)

REFERENCE Genoscope.

AUTHORS Direct Submission

TITLE Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT - Web : www.genoscope.cns.fr)

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetradon>.

FEATURES

Source 1..1070

Location/Qualifiers

/organism="Tetradon nigroviridis"

/mol\_type="genomic DNA"

/db\_xref="taxon:99883"

/clone="054A19"

/clone\_lib="G"

/note="Genoscope sequence ID : COBG054A10LP1-end : T7"

BASE COUNT 244 a 270 c 279 g 270 t

ORIGIN



Query Match 91.6%; Score 17.4; DB 29; Length 1070;  
 Best Local Similarity 94.7%; Pred. No. 1.3e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGCCATCAGTCTCTG 19  
 |||||  
 Db 225 CTGGCCATCAGTCTCTG 243

RESULT 18  
 BG327989/c 1579 bp mRNA linear EST 27-FEB-2001  
 LOCUS 602427610F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:454689 5',  
 DEFINITION mRNA sequence.  
 BG327989  
 ACCESSION BG327989.1 GI:13134427  
 VERSION  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1579)  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: NIH Intramural Sequencing Center  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: LNCM1234 row: d column: 04  
 High quality sequence stop: 197.  
 Location/Qualifiers  
 1. 1579  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4546899"  
 /tissue\_type="adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_15"  
 /note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; CDNA made by oligo-dt priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-CDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 475 a 218 g 113 t  
 ORIGIN 773 c

Query Match 91.6%; Score 17.4; DB 10; Length 1579;  
 Best Local Similarity 94.7%; Pred. No. 1.5e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGCCATCAGTCTCTG 19  
 |||||  
 Db 47 CTGGCCATCAGTCTCTG 29

RESULT 19  
 AK034665/c 1817 bp mRNA linear HTC 05-DEC-2002  
 LOCUS AK034665  
 DEFINITION Mus musculus 12 days embryo embryonic body between diaphragm region  
 and neck CDNA, RIKEN full-length enriched library, clone:943002Z2H21  
 product:unknown EST, full insert sequence.  
 AK034665  
 VERSION AK034665.1 GI:26084119

KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

REFERENCE  
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20493374  
 PUBMED 11042159

REFERENCE  
 3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M.,  
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kaishwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multichipillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

REFERENCE  
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanka, I.,  
 Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, W., Gaasterland, T., Gisi, C., King, B., Kochwa, H.,  
 Kuhl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
 Quackenbush, J., Schmitt, L.M., Straub, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Barah, G., Blake, J., Boffelli, D., Bojunga, N.,  
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Guelinckx, S., Hill, D.,  
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaur, P., Nordone, P.,  
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
 Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,  
 Wyshew-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S.  
 and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 MEDLINE 21085660  
 PUBMED 11217851

REFERENCE  
 5 The PANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

XX [2]  
RN

```
/organism="Homo sapiens"
/mol_type="tRNA"
```

```
1.178
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
```

/clone\_lib="NN0233"  
 /note="Organ: nervous normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."  
 BASE COUNT 45 a 37 c 40 g 56 t  
 ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 178;  
 Best Local Similarity 94.4%; Pred. No. 1.6e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 10 TTGGCCATCAGTGTCTG 19  
 10 TTGGCCATCAGTGTCTG 27

RESULT 22  
 CA912316/c 292 bp mRNA linear EST 27-DEC-2002  
 LOCUS PC803177 Scarlet Runner Bean Suspensor Region Triplex2 Phaseolus  
 DEFINITION coccineus cDNA 3' similar to putative protein; protein id:  
 AC957150. mRNA sequence.  
 AC912316  
 CA912316.1 GI:27399308  
 EST.  
 SOURCE Phaseolus coccineus  
 ORGANISM Phaseolus coccineus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Phaseolus.  
 1 (bases 1 to 292)  
 BuI,A.Q., Le,B.H., Weterings,K., Bi,Y.-P., Choi,J.-S., McElroy,K.E.,  
 Choi,P.S., Harada,J.J., Fischer,R.L. and Goldberg,R.B.  
 Gene Activity in Different Regions of a Post-Fertilization Plant  
 Embryo by EST Analysis  
 Unpublished  
 Contact: Goldberg, R.B.  
 Department of Molecular, Cell, & Developmental Biology  
 University of California, Los Angeles  
 621 Charles E. Young Drive South, Los Angeles, CA 90095-1606, USA  
 Tel: 310 825 3270  
 Fax: 310 825 8201  
 Email: bobg@ucla.edu  
 Seq primer: 5' Triplex  
 POLYA=No.

FEATURES  
 Source location/Qualifiers  
 1..292  
 /organism="Phaseolus coccineus"  
 /mol\_type="mRNA"  
 /cultivar="Hammond's Dwarf Scarlet"  
 /db\_xref="taxon:3886"  
 /date="6-days post-pollination"  
 /clone\_lib="Scarlet Runner Bean Suspensor Region Triplex2"  
 /note="Organ: Suspensor Region of Globular-Stage Embryos;  
 Vector: Triplex2; Site\_1: SfiI; Site\_2: SfiI; Suspensor  
 regions were micro-dissected from globular-stage embryos  
 six days after pollination from greenhouse-grown plants  
 [Weterings et al., Plant Cell 13, 2409-2425 (2001)].  
 Double-stranded cDNA was synthesized from suspensor mRNA  
 using the SMART cDNA Library Construction Kit according to  
 the manufacturer (Clontech). The suspensor cDNA fragments  
 were directionally ligated into the SfiI restriction site  
 of the lambda Triplex2 vector (Clontech) and the  
 recombinant cDNAs were transformed into E. coli XL1-Blue  
 cells (Clontech). Suspensor cDNA plasmid used for  
 directional sequencing were obtained by in vivo excision  
 from the lambda Triplex2 recombinants in E. coli BM25.8  
 cells (Clontech)."

BASE COUNT 95 a 44 c 97 g 56 t  
 ORIGIN

Query Match 86.3%; Score 16.4; DB 14; Length 292;  
 Best Local Similarity 94.4%; Pred. No. 2e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 CTGGGCGATCAGTGTCTG 18  
 172 CTGGGCGATCAGTGTCTG 155

RESULT 23  
 BF328982/c 322 bp mRNA linear EST 22-NOV-2000  
 LOCUS MR4-BN0238-010800-001-e03 BN0238 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BF328982  
 ACCESSION BF328982.1 GI:11299730  
 VERSION  
 KEYWORDS  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 322)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,  
 Goldman,G.H., Carvalho,A.F., Matukuma,A., Bata,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.U. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 JOURNAL MEDITINE  
 PUBLISHED 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL.  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR4ct2=MR4-BN0238-  
 010800-001-e03&t3=2000-08-01&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 28  
 High quality sequence stop: 322.

FEATURES  
 Source location/Qualifiers  
 1..322  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /date="Adult"  
 /clone\_lib="BN0238"  
 /note="Organ: breast normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 113 a 64 c 59 g 86 t  
 ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 322;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 2 TTGGCCATCAGTGTCTG 19  
 2 TTGGCCATCAGTGTCTG 19

Db 224 TGGGCATAGTCTCTG 207

RESULT 24  
AM594895/c  
LOCUS  
DEFINITION AM594895 327 bp mRNA linear EST 22-MAR-2000  
FK26a06.y1 zebrafish fin day1 regeneration Danto reio cDNA 5' similar to SW:IAIP\_CHICK Q90660 INHIBITOR OF APOPTOSIS PROTEIN ; mRNA sequence.

ACCESSION AM594895.1 GI:7282153  
VERSION AM594895.1  
KEYWORDS EST.  
SOURCE Danto reio (zebrafish)  
ORGANISM Danto reio

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 327)  
Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Seepoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.  
Masnu Zebrafish EST Project 1998

TITLE Unpublished  
JOURNAL  
COMMENT Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbrafish@wustl.wustl.edu  
CDNA Library Preparation: Raymond Lee, CDNA Library Arrayed by: Matthew Clark, DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and Resourcenzentrum Primatendbank, Berlin, Germany (web address: www.rzp.de)  
Seq primer: T3 ET from Amersham  
High quality sequence stop: 316.

FEATURES  
source  
1..327  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/sex="mixed male and female"  
/tissue\_type="1 day fin regenerates"  
/lab\_host="E. coli XL0LR"  
/clone\_lib="zebrafish fin day1 regeneration"  
/note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; 1st strand cDNA primed with (GA)10ACTAGTCTGAG(T)18, followed by second strand synthesis, and ligated to 5' adapter (5'-aattcgacagcag-3', 3'-gccggcc-5'). cDNA was cloned directionally (EcoRI/XhoI) into Stratagene Zap express lambda phage arms. Mass invivo excision done to obtain inserts in pBK-CMV phagemid."

BASE COUNT 100 a 67 c 80 g 80 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 327;  
Best Local Similarity 94.4%; Pred. No. 2.1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGGCATAGTCTCT 18  
|||||  
Db 70 CTGGGCATAGTCTCT 53

RESULT 25  
CC469335/c  
LOCUS CC469335 381 bp DNA linear GSS 12-JUN-2003

DEFINITION CH240\_141P20.TV CHORI-240 Bos taurus genomic clone CH240\_141P20, genomic survey sequence.

ACCESSION CC469335  
VERSION CC469335.1 GI:11655567  
KEYWORDS GSS.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
1 (bases 1 to 381)  
Costa, J.N., Motz, M., and Caetano, A.R.  
Brazil's Contribution to End-Sequencing the Bovine BAC Library CHORI-240

JOURNAL Unpublished  
COMMENT Other GSSs: CH240\_141P20.TV  
Contact: Caetano AR  
Department of Biotechnology  
Embrapa Recursos Geneticos e Biotecnologia  
Parque Estacao Biologica, Final Av. W/5 Norte, Brasilia-DF C.P. 02372, 70770-900 Brasil  
Tel: 55 61 448 4778  
Fax: 55 61 340 3658  
Email: acetano@embrapa.br  
Clones are derived from the bovine BAC library CHORI-240 (<http://www.chori.org/bacpac/bovine240.htm>). Bases shown have Phred quality value equal to or higher than 20. Bases with quality value below 20 were masked with 'N'. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources ([http://www.chori.org/bacpac/ordering\\_information.htm](http://www.chori.org/bacpac/ordering_information.htm)). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by Embrapa Recursos Geneticos e Biotecnologia with financing from Conselho Nacional de Desenvolvimento Cientifico e Tecnol6gico (CNPq), Brazil.  
Plate: 141 Row: P Column: 20  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 381.

FEATURES  
source  
1..381  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/strain="bred: Hereford"  
/db\_xref="taxon:9913"  
/clone="CH240\_141P20"  
/sex="Male"  
/cell\_type="Blood"  
/clone\_lib="CHORI-240"  
/note="Vector: pTARBA1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

BASE COUNT 99 a 78 c 82 g 61 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 29; Length 381;  
Best Local Similarity 94.4%; Pred. No. 2.3e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGGCATAGTCTCT 18  
|||||  
Db 219 CTGGTCATAGTCTCT 202

RESULT 26  
CB806515/c  
LOCUS CB806515 386 bp mRNA linear EST 16-MAY-2003  
DEFINITION AMGN:NRDGL-00115-A10-A nrdg1 (10855) Rattus norvegicus cDNA clone nrdg1-00115-a10 5', mRNA sequence.

ACCESSION CB806515  
VERSION CB806515.1 GI:29924649  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 386)  
AUTHORS Angen EST Program.  
TITLE Angen Rat EST Program  
JOURNAL Unpublished  
COMMENT Contact: Dan Fitzpatrick  
Angen, Inc  
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00115 row: a column: 10.  
Location/Qualifiers  
1. 386  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="nr01-00115-a10"  
/tissue\_type="dorsal Root Ganglia"  
/clone\_lib="nr01 (10855)"  
/note="Vector: pSPOR1; Site\_1: SalI; Site\_2: NotI; rat  
dorsal root ganglia"

BASE COUNT 128 a 61 c 82 g 115 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 14; Length 386;  
Best Local Similarity 94.4%; Pred. No. 2.3e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCGCATCAGTCTCT 18  
34 CTGGGCGCATCAGTCTCT 17

RESULT 27  
BY596515/c 422 bp mRNA linear EST 15-DEC-2002  
LOCUS BY596515 RIKEN full-length enriched, adult inner ear Mus musculus  
DEFINITION cDNA clone F330050L08 3, mRNA sequence.  
VERSION BY596515  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 422)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Oosato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,  
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,  
Gojobori, T., Baldarelli, R., Hill, D.P., Bulc, C., Hume, D.A.,  
Quackenbush, J., Schriml, L.M., Kanaphi, A., Matsuda, H., Batalov, S.,  
Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbett,  
L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,  
A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,  
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,  
Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,  
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,  
P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,  
H., Nagashima, T., Numata, K., Okido, T., Pavan, W.U., Pertea, G.,  
Pesole, G., Petrosky, N., Pillai, R., Pontius, J.U., Qi, D.,  
Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,  
B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempile, C.A., Setou,  
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,  
R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,  
Watanabe, Y., Wells, C., Wilming, L.G., Wysshaw-Boris, A., Yanagisawa,  
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,  
Carninci, P., Hayatsu, N., Hirozane-Kikuhawa, T., Komori, H., Nakamura,  
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,  
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,  
K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander

TITLE 'B.S., Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL 22354683  
MEDLINE 12466851  
PubMed

COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 220-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Atsawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane,  
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komori, H., Miyazaki, A.,  
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,  
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,  
M., Waki, K., Watanaki, A., Muramatsu, M. and Hayashizaki, Y. Direct  
Submission

Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by Kirk W. Beisel (Boys Town National  
Research Hospital 555 North 30th Street Omaha, NE 68131 USA ) whose  
assistance we gratefully acknowledge.  
Please visit our web site (http://genome.gsc.riken.go.jp) for  
further details.

FEATURES  
source 1. 422  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="F330050L08"  
/tissue\_type="inner ear"  
/dev\_stage="adult"  
/clone\_lib="RIKEN full-length enriched, adult inner ear"

BASE COUNT 107 a 90 c 118 g 107 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 13; Length 422;  
Best Local Similarity 94.4%; Pred. No. 2.4e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGGGCGCATCAGTCTCTG 19  
173 TGGGCGCATCAGTCTCTG 156

RESULT 28  
A2786165/c 431 bp DNA linear GSS 16-FEB-2001  
LOCUS A2786165 Mouse 10kb plasmid library Mus musculus genomic  
DEFINITION 2M0031F02R Mouse 10kb plasmid library Mus musculus genomic  
clone UGCG2M0031F02 R, genomic survey sequence.  
VERSION A2786165  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 431)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weis,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL Unpublished  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0031 row: F column: 02  
Seq primer: CACACAGGAAACGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 431.  
Location/Qualifiers  
1..431  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG2M0031F02"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCGIM library"  
/note="Vector: PMD42nv, Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptor DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PMD42 (g1|4732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptor mouse DNA was annealed to  
adaptor vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 112 a 95 c 111 g 113 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 28; Length 431;  
Best Local Similarity 94.4%; Pred. No. 2.4e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGCCTCT 18  
|||||  
Db 384 CTGGGCATCAGTGCCTCT 367

RESULT 29  
LOCUS A0706454/c 460 bp DNA linear GSS 07-JUL-1999  
DEFINITION HS\_5554\_A1-G06 SP6E RPCI-11 Human Male BAC Library Homo sapiens  
genomic clone Plate=1130 Col=11 Row=M, genomic survey sequence.  
ACCESSION A0706454  
VERSION A0706454.1 GI:5415880  
KEYWORDS GSS.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 460)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
PUBMED 10449764  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@jong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm)  
or from Research Genetics (info@resgen.com). BAC end Web Server:  
http://www.htsc.washington.edu  
Plate: 1130 row: M column: 11  
Seq primer: SP6  
Class: BAC ends  
High quality sequence stop: 460.  
Location/Qualifiers  
1..460  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=1130 Col=11 Row=M"  
/sex="male"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI methylase. Size selected DNA was cloned into the  
pBAC3.6 vector at EcoRI sites"

BASE COUNT 166 a 71 c 123 g 97 t 3 others  
ORIGIN

Query Match 86.3%; Score 16.4; DB 28; Length 460;  
Best Local Similarity 94.4%; Pred. No. 2.5e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGGGCATCAGTGCCTCTG 19  
|||||  
Db 412 TTGGGCATCAGTGCCTCTG 395

RESULT 30  
LOCUS A2813275 474 bp DNA linear GSS 20-FEB-2001  
DEFINITION 2M0080405R Mouse 10kb plasmid UUCGIM library Mus musculus genomic  
clone UUCG2M0080405 R, genomic survey sequence.  
ACCESSION A2813275  
VERSION A2813275.1 GI:12983183  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 474)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weis,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL  
COMMENT

plasmid inserts  
 unpublished  
 Contact: Robert B. Weiss  
 University of Utah  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: dduh@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Seq primer: CACACAGGAACACCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 474.

## FEATURES

## source

1. .474  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="U05C2M080A05"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid U05C2M library"  
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN

78 a 143 c 136 g 117 t

Query Match 86.3%; Score 16.4; DB 28; Length 474;  
 Best Local Similarity 94.4%; Pred. No. 2.5e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCT 18  
 Db 337 CTGGGCCATCAGTCTCT 354

## RESULT 31

LOCUS B1473885 477 bp mRNA linear EST 27-AUG-2001  
 DEFINITION fp48b06.y3 zebrafish gridded kidney Danio rerio cDNA clone  
 IMAGE:4759643 5' similar to SW:1AP\_CHICK 090660 INHIBITOR OF  
 APOPTOSIS PROTEIN ;, mRNA sequence.

ACCESSION B1473885  
 VERSION B1473885.1 GI:15302437  
 KEYWORDS EST.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
 1 (bases 1 to 477)  
 AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy  
 S., Hillier,L., Kucaba,T., Martin,T., Beck,C., Wylie,T., Underwood

TITLE  
JOURNAL  
COMMENT

Unpublished  
 Contact: Stephen L. Johnson  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: zbratfish@wustl.edu  
 CDNA Library Preparation: Leonard Ira Zon DNA Sequencing by:  
 Washington University Genome Sequencing Center Clone distribution:  
 Genome Systems, St. Louis, Missouri (web address:  
 www.genomesystems.com) (email contact: info@genomesystems.com) and  
 Research Genetics, Huntsville, Alabama (web address: www.resgen.com)  
 (email contact: info@resgen.com) and  
 Ressourcenzentrum für Primatendatbank, Berlin, Germany (web address:  
 www.rzp.de)  
 Possible reversed clone: similarity on wrong strand  
 High quality sequence stop: 294.

FEATURES  
source

1. .477  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"  
 /clone="IMAGE:4759643"  
 /sex="mixed"  
 /tissue\_type="kidney pooled from 300 wild type adults"  
 /lab\_host="XLOLR"  
 /clone\_lib="zebrafish gridded kidney"  
 /note="Organ: kidney; Vector: pBK-CMV; Site 1: ECORI;  
 Site 2: XhoI; Oligo dT cDNA library constructed from mRNA  
 pooled from pooled kidney tissue from 300 adult  
 zebrafish."

BASE COUNT  
ORIGIN

114 a 109 c 96 g 158 t

Query Match 86.3%; Score 16.4; DB 12; Length 477;  
 Best Local Similarity 94.4%; Pred. No. 2.5e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCT 18  
 Db 289 CTGGGCCATCAGTCTCT 306

## RESULT 32

LOCUS BG728994 522 bp mRNA linear EST 09-MAY-2001  
 DEFINITION fp06g01.x1 zebrafish gridded kidney Danio rerio cDNA clone  
 IMAGE:4728769 3' similar to SW:1AP\_CHICK 090660 INHIBITOR OF  
 APOPTOSIS PROTEIN ;, mRNA sequence.

ACCESSION BG728994  
 VERSION BG728994.1 GI:14014069  
 KEYWORDS EST.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
 1 (bases 1 to 522)  
 AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy  
 S., Hillier,L., Kucaba,T., Martin,T., Beck,C., Wylie,T., Underwood  
 S., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,  
 Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,B.,  
 Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.  
 and Wilson,R.  
 WASHU Zebrafish EST Project 1998  
 Unpublished  
 Contact: Stephen L. Johnson

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: zbratish@watson.wustl.edu  
CDNA Library Preparation: Leonard Ira Zon DNA Sequencing by:  
Washington University Genome Sequencing Center Clone Distribution:  
Genome Systems, St. Louis, Missouri (web address:  
www.genomesystems.com) (email contact: info@genomesystems.com) and  
Research Genetics, Huntsville, Alabama (web address: www.resgen.com  
) (email contact: info@resgen.com) and  
ReSourceCenter/PrimateDataBank, Berlin, Germany (web address:  
www.rzpd.de)

Seq primer: T7 from Gibco  
High quality sequence stop: 361.

FEATURES  
source  
Location/Qualifiers

1..528  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:4728769"  
/sex="mixed"  
/tissue\_type="kidney pooled from 300 wild type adults"  
/lab\_host="XLOLR"  
/clone\_lib="zebrafish gridded kidney"  
/note="Organ: kidney; Vector: PBK-CMV; Site 1: EcoRI;  
Site 2: XhoI; Oligo dt cDNA library constructed from mRNA  
pooled from pooled kidney tissue from 300 adult  
zebrafish."

BASE COUNT 133 a 123 c 106 g 160 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 522;  
Best Local Similarity 94.4%; Pred. No. 2.7e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 CTGGGCGCATCAGTCTCT 18  
|||||  
254 CTGGGCGCATCAGTCTCT 271

RESULT 33 528 bp mRNA linear EST 04-JUN-2001  
BG891847  
LOCUS fpa5e01.x1 zebrafish gridded kidney Danio rerio cDNA clone  
IMAGE:4759752.3' similar to SW:IAIP\_CHICK Q09660 INHIBITOR OF  
APOPTOSIS PROTEIN ; mRNA sequence.

ACCESSION BG891847  
VERSION BG891847.1 GI:14286457  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 528)  
AUTHORS Clark.M., Johnson.S.L., Lehrach.H., Lee.R., Li.F., Marra.M., Eddy  
, S., Hillier.L., Kucaba.T., Martin.J., Beck.C., Wylie.T., Underwood  
, K., Steptoe.M., Theising.B., Allen.M., Bowers.Y., Person.B.,  
Swaller.T., Gibbons.M., Pape.D., Harvey.N., Schurk.R., Ritter.E.,  
Kohn.S., Shin.T., Jackson.Y., Cardenas.M., McCann.R., Waterston.R.  
and Wilson.R.  
Title Wabnu Zebrafish EST Project 1998  
Unpublished

JOURNAL  
COMMENT Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: zbratish@watson.wustl.edu  
CDNA Library Preparation: Leonard Ira Zon DNA Sequencing by:  
Washington University Genome Sequencing Center Clone Distribution:  
Genome Systems, St. Louis, Missouri (web address:

www.genomesystems.com) (email contact: info@genomesystems.com) and  
Research Genetics, Huntsville, Alabama (web address: www.resgen.com  
) (email contact: info@resgen.com) and  
ReSourceCenter/PrimateDataBank, Berlin, Germany (web address:  
www.rzpd.de)  
High quality sequence stop: 374.

FEATURES  
source  
Location/Qualifiers

1..528  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:4759752"  
/sex="mixed"  
/tissue\_type="kidney pooled from 300 wild type adults"  
/lab\_host="XLOLR"  
/clone\_lib="zebrafish gridded kidney"  
/note="Organ: kidney; Vector: PBK-CMV; Site 1: EcoRI;  
Site 2: XhoI; Oligo dt cDNA library constructed from mRNA  
pooled from pooled kidney tissue from 300 adult  
zebrafish."

BASE COUNT 132 a 126 c 108 g 162 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 528;  
Best Local Similarity 94.4%; Pred. No. 2.7e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 CTGGGCGCATCAGTCTCT 18  
|||||  
253 CTGGGCGCATCAGTCTCT 270

RESULT 34 536 bp mRNA linear EST 26-JUN-2002  
BO616929  
LOCUS fab02a02.x1 Sugano SJD adult male Danio rerio cDNA clone  
IMAGE:6034059.3' similar to SW:IAIP2\_HUMAN Q13490 INHIBITOR OF  
APOPTOSIS PROTEIN 2 ; mRNA sequence.

ACCESSION BO616929  
VERSION BO616929.1 GI:21606598  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 536)  
AUTHORS Clark.M., Johnson.S.L., Lehrach.H., Lee.R., Li.F., Marra.M., Eddy  
, S., Hillier.L., Kucaba.T., Martin.J., Beck.C., Wylie.T., Underwood  
, K., Steptoe.M., Theising.B., Allen.M., Bowers.Y., Person.B.,  
Swaller.T., Gibbons.M., Pape.D., Harvey.N., Schurk.R., Ritter.E.,  
Kohn.S., Shin.T., Jackson.Y., Cardenas.M., McCann.R., Waterston.R.  
and Wilson.R.  
Title Wabnu Zebrafish EST Project 1998  
Unpublished

JOURNAL  
COMMENT Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: zbratish@watson.wustl.edu  
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA  
Sequencing by: Washington University Genome Sequencing Center Clone  
Distribution information can be found through the I.M.A.G.E.  
Consortium/LIML, send email to: info@image.liml.gov  
Seq primer: T7 from Gibco.

FEATURES  
source  
Location/Qualifiers  
1..536  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:6034059"  
/sex="male"



```
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/clone_lib="Sugano SJD adult male"
/note="Vector: pMR18-Flu3; Site 1: DraIII (CACCATG);
Site 2: DraIII (CACTGTC); 1st strand cDNA was primed
with an oligo(dT) primer [ATGGGCGCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GTGGCGCTACTGG], digested and cloned into distinct DraIII
sites of the pMR18-Flu3 vector (5' site CACATGTC, 3' site
CACCATGTC). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed and donated by Dr. Sumio
Sugano (University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTTAAGAAGCTGCG and 3' end primer
CGACCTCGAGCTCGAGCACA."
```

BASE COUNT 132 a 126 c 107 g 171 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 13; Length 536;  
Best Local Similarity 94.4%; Pred. No. 2.7e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCT 18  
|||||  
264 CTGGGCCATCAGTCTCT 261

Db

RESULT 35  
BQ074936/c  
LOCUS 574 bp mRNA linear EST 02-APR-2002  
DEFINITION fz24f09.y1 Gong zebrafish testis Danio rerio cDNA 5' similar to  
SW-1A2\_HUMAN Q13490 INHIBITOR OF APOPTOSIS PROTEIN 2 ;, mRNA  
sequence.  
ACCESSION BQ074936  
VERSION BQ074936.1 GI:19903973  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio

REFERENCE  
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy  
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood  
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,  
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler,E.,  
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.  
and Wilson,R.

TITLE Wabhu Zebrafish EST Project 1998  
JOURNAL Unpublished  
COMMENT Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbrafish@wustl.edu  
The library was constructed by Dr. Z. Gong. DNA Sequencing by:  
Washington University Genome Sequencing Center St. Louis. Please  
contact Zhiyuan Gong for further information on this library  
(National University of Singapore, Department of Biological  
Sciences, Lower Kent Ridge Road, Singapore 119260).  
Seq primer: T3 ET from Amersham  
High quality sequence stop: 424.  
Location/Qualifiers  
1..574  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/sex="male"  
/dev\_stage="4-5 month"  
/lab\_host="DH10B (phage-resistant)"

FEATURES  
source

```
/clone_lib="Gong zebrafish testis"  
/note="Organ: testis (pooled); Vector: pBluescript SK-;  
Site 1: XhoI; Site 2: EcoRI; Poly A+ RNA was isolated from  
the testes of 31 male adult zebrafish (4-5 month old).  
cDNAs were made using oligo-dT primers and inserted into  
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo  
mass-excised to pBluescript SK- following the Washington  
University protocol  
(http://genome.wustl.edu/lambda_protocol.shtml).  
Please contact Zhiyuan Gong for further information on  
this library (National University of Singapore,  
Department of Biological Sciences, Lower Kent Ridge Road,  
Singapore 119260)."
```

BASE COUNT 179 a 106 c 154 g 135 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 13; Length 574;  
Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCT 18  
|||||  
430 CTGGGCCATCAGTCTCT 413

Db

RESULT 36  
B1670955  
LOCUS 591 bp mRNA linear EST 13-FEB-2002  
DEFINITION f58f06.x1 Zebrafish adult retina cDNA Danio rerio cDNA clone  
IMAGE:4786307 3' similar to SW:1AP\_CHICK Q90660 INHIBITOR OF  
APOPTOSIS PROTEIN ;, mRNA sequence.  
ACCESSION B1670955  
VERSION B1670955.1 GI:15586339  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio

REFERENCE  
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy  
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood  
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,  
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler,E.,  
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.  
and Wilson,R.

TITLE Wabhu Zebrafish EST Project 1998  
JOURNAL Unpublished  
COMMENT Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbrafish@wustl.edu  
Library constructed by: Chandra Tucker and Gregory Niemi DNA  
Sequencing by: Washington University Genome Sequencing Center Clone  
Distribution: RessourcenzentrumPrimarDatbank, Berlin, Germany  
(web address: www.rzpd.de)  
Seq primer: T7 from Gibco  
High quality sequence stop: 389.  
Location/Qualifiers  
1..591  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/strain="wild-type"  
/db\_xref="taxon:7955"  
/clone="IMAGE:4786307"  
/sex="mixed"  
/dev\_stage="1-2 years"  
/lab\_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')"  
/clone\_lib="Zebrafish adult retina cDNA"  
/note="Vector: lambda ZAP II (pBluescript SK-); Site 1:

FEATURES  
source

ECORI; Site 2: SalI; This zebrafish library was constructed by Dr. Susan E. Brockhoff (email: sbrockhoff@u.washington.edu) RZPD library number: 760"

BASE COUNT 142 a 148 c 113 g 188 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 591;  
Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCT 18  
Db 263 CTGGGCCATCAGTCTCT 280

RESULT 37 596 bp mRNA linear EST 14-NOV-2002  
CA500347/C LOCUS

DEFINITION WHE4018.D05.G10ZT Wheat meiotic anther cDNA library Triticum aestivum cDNA clone WHE4018.D05.G10, mRNA sequence.

ACCESSION CA500347 GI:24991307  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Iazo, G.R., Pham, J., Rausch, C.J., Sutton, T., Woo, J., and Wilson, C.

TITLE The structure and function of the expressed portion of the wheat genomes - Meiotic anther cDNA library

JOURNAL Unpublished  
COMMENT Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderson@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
Seq primer: T7 primer.

FEATURES  
source

1..596 Location/Qualifiers

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE4018.D05.G10"

/tissue\_type="Anther"

/dev\_stage="Meiotic stages pre-meiosis-metaphase I"

/lab\_host="E. coli DH10B"

/clone\_lib="Wheat meiotic anther cDNA library"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants were grown in a glasshouse. Anther meiotic stage was determined by removing anthers from individual primary florets. One anther was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two anthers were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared. cDNA synthesised, and directionally ligated into pSPORT1 by Tim Sutton in the P Langridge Lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5Kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson Lab (all other authors)."

BASE COUNT 200 a 119 c 164 g 113 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 14; Length 596;

Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CTGGGCCATCAGTCTCT 18  
Db 552 CTGGGCCATCAGTCTCT 535

RESULT 38 602 bp mRNA linear EST 07-JUN-2001  
A1496684/C LOCUS

DEFINITION FB48e09.y1 Zebrafish Washu MPMG EST Danio rerio cDNA clone IMAGE:375144.5' similar to SW:1AP CHICK G90650 INHIBITOR OF APOPTOSIS PROTEIN 1, mRNA sequence.

ACCESSION A1496684 GI:4397687  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, D., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

TITLE Washu Zebrafish EST Project 1998  
JOURNAL Unpublished  
COMMENT Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbrfish@wustl.edu

CDNA Library Preparation: Matthew Clark, cDNA Library Arrayed by: Matthew Clark, DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and Resourcenzentrum Primatdatenbank, Berlin, Germany (web address: www.rzpd.de)

Possible reversed clone: similarity on wrong strand  
Seq primer: T3 ET from Amersham  
High quality sequence stop: 462  
POLYA-No.

FEATURES  
source

1..602 Location/Qualifiers

/organism="Danio rerio"

/mol\_type="mRNA"

/db\_xref="taxon:7955"

/clone="IMAGE:375144"

/sex="mixed"

/tissue\_type="26 somite embryos, adult livers, shield stage embryos"

/lab\_host="XLI-blue MRP"

/clone\_lib="Zebrafish Washu MPMG EST"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer 15'GACAGTTCATGATCGGAGCGCGCCCTTTTCTTTT3'; double-stranded cDNA was ligated into Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach Lab), ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a

single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

```

BASE COUNT      179 a      123 c      154 g      146 t
ORIGIN
Query Match      86.3%; Score 16.4; DB 9; Length 602;
Best Local Similarity 94.4%; Pred. No. 2.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CTGGGCCATCAGTCTCT 18
      |||||
Db      374 CTGGGCCATCAGTCTCT 357

RESULT 39
LOCUS      B0074707
DEFINITION B0074707      627 bp      mRNA      linear      EST 02-APR-2002
            fz24f09.x1 Gong zebrafish testis Danio rerio cDNA 3' similar to
            SW:JAP_CHICK Q90660 INHIBITOR OF APOPTOSIS PROTEIN 1; mRNA
            sequence.
ACCESSION  B0074707
VERSION    B0074707
KEYWORDS   Danio rerio (zebrafish)
SOURCE     EST.
ORGANISM   Danio rerio

REFERENCE
AUTHORS    Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            R., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuck,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
            Washu Zebrafish EST Project 1998
TITLE      Unpublished
JOURNAL
COMMENT    Other ESTs: fz24f09.y1
            Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbratfish@wustl.edu
            The library was constructed by Dr. Z. Gong. DNA sequencing by:
            Washington University Genome Sequencing Center St. Louis. Please
            contact Zhiyuan Gong for further information on this library
            (National University of Singapore, Department of Biological
            Sciences, Lower Kent Ridge Road, Singapore 119260).
            Seq primer: T7 from Gibco
            High quality sequence stop: 473.
            Location/Qualifiers
            1..627
               /organism="Danio rerio"
               /mol_type="mRNA"
               /db_xref="taxon:7955"
               /sex="male"
               /dev_stage="4-5 month"
               /lab_host="DH10B (phage-resistant)"
               /clone_1lb="Gong zebrafish testis"
               /note="Organ: testis (pooled); Vector: Bluescript SK-;
               Site 1: XhoI; Site 2: EcoRI; Poly A+ RNA was isolated from
               the testes of 31 male adult zebrafish (4-5 month old).
               cDNAs were made using oligo-dT primers and inserted into
               lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
               mass-excised to Bluescript SK- following the Washington
               University protocol
               (http://genome.wustl.edu/est/lambda_protocol.shtml).
               Please contact Zhiyuan Gong for further information on
               this library (National University of Singapore,

```

```

BASE COUNT      149 a      156 c      119 g      203 t
ORIGIN
Query Match      86.3%; Score 16.4; DB 13; Length 627;
Best Local Similarity 94.4%; Pred. No. 2.9e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CTGGGCCATCAGTCTCT 18
      |||||
Db      265 CTGGGCCATCAGTCTCT 282

RESULT 40
LOCUS      BE557791
DEFINITION BE557791      672 bp      mRNA      linear      EST 30-AUG-2000
            f110g10.y1 Zebrafish Research Genetics C32 fin Danio rerio cDNA 5'
            similar to SW:JAP2_HUMAN Q13490 INHIBITOR OF APOPTOSIS PROTEIN 2 ;
            mRNA sequence.
ACCESSION  BE557791
VERSION    BE557791
KEYWORDS   Danio rerio (zebrafish)
SOURCE     EST.
ORGANISM   Danio rerio

REFERENCE
AUTHORS    Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            R., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuck,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
            Washu Zebrafish EST Project 1998
TITLE      Unpublished
JOURNAL
COMMENT    Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbratfish@wustl.edu
            cDNA Library Preparation: Ning Wu. cDNA library Arrayed by:
            Research Genetics. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Research Genetics web
            address: http://www.researchgenetics.com/
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 465.
            Location/Qualifiers
            1..672
               /organism="Danio rerio"
               /mol_type="mRNA"
               /db_xref="taxon:7955"
               /issue_type="Fin"
               /lab_host="GeneHogs (HS996, a phage-resistant isolate of
               DH10B)"
               /clone_1lb="Zebrafish Research Genetics C32 fin"
               /note="Vector: pRT73D-Pac with a modified polylinker;
               Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was prepared
               from zebrafish(C32) fin, and was then primed with a Not I
               - oligo(dT) primer. Double-stranded cDNA was ligated to
               Eco RI adaptors (Pharmacia), digested with Not I and
               cloned into the Not I and Eco RI sites of the modified
               pRT73 vector. Library is non-normalized. Library was
               constructed by Ning Wu. NOTE: This clone is available
               royally-free through LBNL; contact the IMGS Consortium
               (info.lbnl.gov) for further information"

BASE COUNT      159 a      181 c      127 g      205 t
ORIGIN
Query Match      86.3%; Score 16.4; DB 10; Length 672;
Best Local Similarity 94.4%; Pred. No. 3e+03;

```

Matches 17; Conservativ 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CTGGGCCATCAGTCTCT 18

Db 220 CTGGGCCATCAGTCTCT 237

Search completed: January 13, 2004, 16:33:23  
Job time : 1196.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 13, 2004, 17:32:17 ; Search time 126.5 Seconds  
(without alignments)  
521.578 Million cell updates/sec

Title: US-09-719-737-18

Perfect score: 19  
Sequence: 1 ctgggcatcagtcctcgc 19

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2276164 seqs, 1736306516 residues

Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	591	13	US-10-029-386-9548
2	19	100.0	1065	10	US-09-922-895-2
3	19	100.0	1068	13	US-10-276-950-2
4	19	100.0	1068	13	US-09-826-509-474
5	19	100.0	1071	16	US-10-293-050-6
6	19	100.0	1116	13	US-10-283-028-5
7	19	100.0	1193	13	US-10-283-028-3
8	19	100.0	1201	15	US-10-225-567A-63
9	19	100.0	1689	13	US-09-931-381A-15
10	19	100.0	1689	13	US-10-283-028-1
11	19	100.0	1717	10	US-09-964-824A-100
12	19	100.0	1915	14	US-10-106-623-3
13	19	100.0	3426	14	US-10-001-835-29
14	17.4	91.6	549	9	US-09-864-761-9611
15	17.4	91.6	563	13	US-10-029-386-152

C 16	17.4	91.6	1516	10	US-09-745-605-2	Sequence 2, App1
C 17	16	84.2	824	13	US-10-027-632-136447	Sequence 136447,
C 18	16	84.2	824	13	US-10-027-632-136511	Sequence 136511,
C 19	16	84.2	824	14	US-10-027-632-136447	Sequence 136447,
C 20	16	84.2	824	14	US-10-027-632-136511	Sequence 136511,
C 21	15.8	83.2	145	13	US-10-188-619-14	Sequence 14, App1
C 22	15.8	83.2	145	11	US-09-747-377-447	Sequence 447, App
C 23	15.8	83.2	145	15	US-10-105-613-447	Sequence 583, App
C 24	15.8	83.2	422	10	US-09-920-300A-583	Sequence 583, App
C 25	15.8	83.2	422	13	US-10-099-926-583	Sequence 583, App
C 26	15.8	83.2	422	14	US-10-033-528-583	Sequence 583, App
C 27	15.8	83.2	492	13	US-10-027-632-281562	Sequence 281562,
C 28	15.8	83.2	492	13	US-10-027-632-281563	Sequence 281563,
C 29	15.8	83.2	492	14	US-10-027-632-281562	Sequence 281562,
C 30	15.8	83.2	492	14	US-10-027-632-281563	Sequence 281563,
C 31	15.8	83.2	652	13	US-10-027-632-81346	Sequence 81346, A
C 32	15.8	83.2	652	13	US-10-027-632-81347	Sequence 81347, A
C 33	15.8	83.2	652	13	US-10-027-632-109771	Sequence 109771,
C 34	15.8	83.2	652	13	US-10-027-632-109772	Sequence 109772,
C 35	15.8	83.2	652	13	US-10-027-632-301812	Sequence 301812,
C 36	15.8	83.2	652	13	US-10-027-632-301813	Sequence 301813,
C 37	15.8	83.2	652	14	US-10-027-632-81346	Sequence 81346, A
C 38	15.8	83.2	652	14	US-10-027-632-81347	Sequence 81347, A
C 39	15.8	83.2	652	14	US-10-027-632-109771	Sequence 109771,
C 40	15.8	83.2	652	14	US-10-027-632-109772	Sequence 109772,
C 41	15.8	83.2	652	14	US-10-027-632-301812	Sequence 301812,
C 42	15.8	83.2	652	14	US-10-027-632-301813	Sequence 301813,
C 43	15.8	83.2	1025	13	US-10-027-632-262488	Sequence 262488,
C 44	15.8	83.2	1025	13	US-10-027-632-262489	Sequence 262489,
C 45	15.8	83.2	1025	14	US-10-027-632-262488	Sequence 262488,

#### ALIGNMENTS

RESULT 1  
US-10-029-386-9548/C  
; Sequence 9548, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 9548  
; LENGTH: 591  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR3.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8  
; OTHER INFORMATION: SWISSPROT HIT: P51677, EVALUATE 1.00e-102  
; OTHER INFORMATION: NT HIT: g11449889, EVALUATE 0.00e+00  
; OTHER INFORMATION: EST\_HUMAN HIT: BG182330.1, EVALUATE 0.00e+00  
US-10-029-386-9548

Query Match 100.0%; Score 19; DB 13; Length 591;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGGGCATCAGTCTCTG 19  
|||||

Db 59 CTGGGCATCAGTGCTCTG 41

## RESULT 2

US-09-922-895-2/c  
; Sequence 2, Application US/09922895  
; Publication No. US20020192214A1  
; GENERAL INFORMATION:  
; APPLICANT: DAUGHERTY, BRUCE L.  
; DEMARTINO, JULIE A.  
; STICLIANO, SALVATORE J.  
; SPRINGER, MARTIN J.  
; TITLE OF INVENTION: EOSINOPHIL EOTAXIN RECEPTOR  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Merck & Co., Inc.  
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065-0900  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/922,895  
; FILING DATE: 06-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/847,296  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 60/017,113  
; FILING DATE: 26-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Thies, J. Eric  
; REGISTRATION NUMBER: 35,382  
; REFERENCE/DOCKET NUMBER: 19634Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 908-594-4720  
; TELEFAX: 908-594-3904  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1065 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-922-895-2

Query Match 100.0%; Score 19; DB 10; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGCTCTG 19  
Db 105 CTGGGCATCAGTGCTCTG 87

RESULT 3  
US-10-276-950-2/c  
; Sequence 2, Application US/10276950  
; Publication No. US20030157639A1  
; GENERAL INFORMATION:  
; APPLICANT: GLAXO GROUP LIMITED  
; TITLE OF INVENTION: PROTEIN AND ASSAY  
; FILE REFERENCE: PG3810  
; CURRENT APPLICATION NUMBER: US/10/276,950  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: GB 0013345.4

; PRIOR FILING DATE: 2000-06-01  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1068  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1068)  
US-10-276-950-2

Query Match 100.0%; Score 19; DB 13; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGCTCTG 19  
Db 105 CTGGGCATCAGTGCTCTG 87

## RESULT 4

US-09-826-509-474/c  
; Sequence 474, Application US/09826509  
; Publication No. US20030204073A1  
; GENERAL INFORMATION:  
; APPLICANT: Lehmann-Bruinema, Karin  
; APPLICANT: Liaw, Chen W.  
; TITLE OF INVENTION: NO. US20030204073A1-Endogenous, Constitutively Activated Known G  
; TITLE OF INVENTION: Protein-Coupled Receptors  
; FILE REFERENCE: AREN-207  
; CURRENT APPLICATION NUMBER: US/09/826,509  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: 60/195,747  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: 09/170,496  
; PRIOR FILING DATE: 1998-10-13  
; NUMBER OF SEQ ID NOS: 589  
; SOFTWARE: PatentIn Version 2.1  
; SEQ ID NO 474  
; LENGTH: 1068  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-826-509-474

Query Match 100.0%; Score 19; DB 13; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGCTCTG 19  
Db 105 CTGGGCATCAGTGCTCTG 87

## RESULT 5

US-10-293-050-6/c  
; Sequence 6, Application US/10293050  
; Publication No. US20030130494A1  
; GENERAL INFORMATION:  
; APPLICANT: Dairaghi, Daniel J.  
; APPLICANT: Hara, Takahiko  
; APPLICANT: Miyajima, Atsushi  
; APPLICANT: Schall, Thomas J.  
; APPLICANT: Wang, Wei  
; APPLICANT: Yoshimura, Akihiko  
; TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS  
; FILE REFERENCE: DX0506B US  
; CURRENT APPLICATION NUMBER: US/10/293,050  
; CURRENT FILING DATE: 2003-02-11  
; PRIOR APPLICATION NUMBER: US 08/567,882  
; PRIOR FILING DATE: 1995-12-08  
; NUMBER OF SEQ ID NOS: 9

SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6  
LENGTH: 1071  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1071)  
OTHER INFORMATION:  
US-10-293-050-6

Query Match 100.0%; Score 19; DB 16; Length 1071;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
DB 105 CTGGGCCATCAGTCTCTG 87

## RESULT 6

US-10-283-028-5/C  
Sequence 5, Application US/10283028  
Publication No. US20030143684A1

## GENERAL INFORMATION:

APPLICANT: Gerard, Craig J.  
Mackay, Charles R.  
Ponath, Paul D.  
Post, Theodore W.  
Qin, Shixin

TITLE OF INVENTION: G PROTEIN-COUPLED RECEPTOR GENE CCR3 AND  
ANTAGONISTS THEREOF

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/283,028

FILING DATE: 28-Oct-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,565

FILING DATE: 30-SEP-1996  
APPLICATION NUMBER: PCT/US96/00608

FILING DATE: 19-JAN-1996  
APPLICATION NUMBER: US 08/375,199

FILING DATE: 19-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.

REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: LKS94-05A2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1116 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-10-283-028-5

Query Match 100.0%; Score 19; DB 13; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
DB 119 CTGGGCCATCAGTCTCTG 101

## RESULT 7

US-10-283-028-3/C  
Sequence 3, Application US/10283028  
Publication No. US20030143684A1

## GENERAL INFORMATION:

APPLICANT: Gerard, Craig J.  
Mackay, Charles R.  
Ponath, Paul D.  
Post, Theodore W.  
Qin, Shixin

TITLE OF INVENTION: G PROTEIN-COUPLED RECEPTOR GENE CCR3 AND  
ANTAGONISTS THEREOF

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/283,028

FILING DATE: 28-Oct-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,565

FILING DATE: 30-SEP-1996  
APPLICATION NUMBER: PCT/US96/00608

FILING DATE: 19-JAN-1996  
APPLICATION NUMBER: US 08/375,199

FILING DATE: 19-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.

REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: LKS94-05A2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1193 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 92..1156

SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-283-028-3

Query Match 100.0%; Score 19; DB 13; Length 1193;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
|||||  
DB 196 CTGGGCCATCAGTCTCTG 178

## RESULT 8

US-10-225-567A-63/c  
; Sequence 63, Application US/10225567A  
; Publication No. US20030113798A1  
; GENERAL INFORMATION:  
; APPLICANT: Lifespan Biosciences  
; APPLICANT: Brown, Joseph P.  
; APPLICANT: Burner, Glenn C.  
; APPLICANT: Roush, Christine L.  
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS  
; FILE REFERENCE: 1920-4-4  
; CURRENT APPLICATION NUMBER: US/10/225,567A  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/257,144  
; NUMBER OF SEQ ID NOS: 2292  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 63  
; LENGTH: 1201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-225-567A-63

Query Match 100.0%; Score 19; DB 15; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGCTCTG 19  
|||||

DB 136 CTGGGCCATCAGTGCTCTG 118  
|||||

## RESULT 9

US-09-931-381A-15/c  
; Sequence 15, Application US/09931381A  
; Patent No. US20020137107A1  
; GENERAL INFORMATION:  
; APPLICANT: Butcher, Eugene C.  
; APPLICANT: Kunkel, Eric J.  
; APPLICANT: Pan, Junliang  
; APPLICANT: Soler-Ferran, Dulce  
; TITLE OF INVENTION: Method for Identifying Agents Which  
; Modulate Chemokine "MCC"-Induced Functions of CCR3 and/or  
; TITLE OF INVENTION: CCR10  
; FILE REFERENCE: 1855.2010-003  
; CURRENT APPLICATION NUMBER: US/09/931,381A  
; CURRENT FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: U.S. 09/638,914  
; PRIOR FILING DATE: 2000-08-15  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 15  
; LENGTH: 1689  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (181)...(1248)  
; NAME/KEY: misc feature  
; LOCATION: (1291)...(1291)  
; OTHER INFORMATION: n = A, T, C or G  
US-09-931-381A-15

Query Match 100.0%; Score 19; DB 10; Length 1689;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGCTCTG 19  
|||||

DB 285 CTGGGCCATCAGTGCTCTG 267  
|||||

## RESULT 10

US-10-283-028-1/c  
; Sequence 1, Application US/10283028  
; Publication No. US20030143684A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerard, Craig J.  
; APPLICANT: Gerard, No. US20030143684A1a P.  
; APPLICANT: Mackay, Charles R.  
; APPLICANT: Ponach, Paul D.  
; APPLICANT: Post, Theodore W.  
; TITLE OF INVENTION: G PROTEIN-COUPLED RECEPTOR GENE CCR3 AND  
; ANTAGONISTS THEREOF  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Millitia Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/283,028  
; FILING DATE: 28-Oct-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/720,565  
; FILING DATE: 30-SEP-1996  
; APPLICATION NUMBER: PCT/US96/00608  
; FILING DATE: 19-JAN-1996  
; APPLICATION NUMBER: US 08/375,199  
; FILING DATE: 19-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: LKS94-05A2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1689 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-283-028-1

Query Match 100.0%; Score 19; DB 13; Length 1689;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGCTCTG 19  
|||||

DB 285 CTGGGCCATCAGTGCTCTG 267  
|||||

## RESULT 11

US-09-964-824A-100/c  
; Sequence 100, Application US/09964824A  
; Patent No. US20020102531A1  
; GENERAL INFORMATION:  
; APPLICANT: Horrigan, Stephen  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE REFERENCE: 689290-73



CURRENT APPLICATION NUMBER: US/09/964,824A  
CURRENT FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: US/60/236,033  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US/60/236,032  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US/60/236,028  
PRIOR FILING DATE: 2000-09-28  
NUMBER OF SEQ ID NOS: 583  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 100  
LENGTH: 1717  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-964-824A-100

Query Match 100.0%; Score 19; DB 10; Length 1717;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTCTCTG 19  
Db 309 CTGGGCATCAGTCTCTG 291

RESULT 12  
US-10-106-623-3/C  
Sequence 3, Application US/10106623  
Publication No. US20020150888A1  
GENERAL INFORMATION:  
APPLICANT: Gray, Patrick W.  
Schweickart, Vicky L.  
REPORT, Carol J.  
TITLE OF INVENTION: Chemokine Receptor Materials and Methods  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/106,623  
FILING DATE: 26-Mar-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/771,276  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: No. US20020150888A1and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/33670  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1915 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 362..1426  
FEATURE:  
NAME/KEY: misc\_feature

OTHER INFORMATION: /= "88-2B polynucleotide and amino acid sequences"  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-106-623-3

Query Match 100.0%; Score 19; DB 14; Length 1915;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTCTCTG 19  
Db 466 CTGGGCATCAGTCTCTG 448

RESULT 13  
US-10-001-835-29/C  
Sequence 29, Application US/10001835  
Publication No. US20020160387A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Recipon, Hervé  
APPLICANT: Calferkey, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Proteins  
REFERENCE: DEX-0277  
CURRENT APPLICATION NUMBER: US/10/001,835  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/249,997  
PRIOR FILING DATE: 2000-11-20  
NUMBER OF SEQ ID NOS: 228  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 29  
LENGTH: 3426  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-001-835-29

Query Match 100.0%; Score 19; DB 14; Length 3426;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTCTCTG 19  
Db 234 CTGGGCATCAGTCTCTG 216

RESULT 14  
US-09-864-761-9611/C  
Sequence 9611, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wenheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmlca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666



```
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136447
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-136447
```

```
Query Match      84.2%; Score 16; DB 13; Length 824;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 GGCATCAGTGTCTG 19
         |||||
Db       342 GGCATCAGTGTCTG 357
```

```
RESULT 18
US-10-027-632-165611
; Sequence 165611, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 165611
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-165611
```

```
Query Match      84.2%; Score 16; DB 13; Length 824;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 GGCATCAGTGTCTG 19
         |||||
Db       342 GGCATCAGTGTCTG 357
```

```
RESULT 19
US-10-027-632-136447
; Sequence 136447, Application US/10027632
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
```

```
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136447
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-136447
```

```
Query Match      84.2%; Score 16; DB 14; Length 824;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 GGCATCAGTGTCTG 19
         |||||
Db       342 GGCATCAGTGTCTG 357
```

```
RESULT 20
US-10-027-632-165611
; Sequence 165611, Application US/10027632
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 165611
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-165611
```

```
Query Match      84.2%; Score 16; DB 14; Length 824;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 GGCATCAGTGTCTG 19
         |||||
Db       342 GGCATCAGTGTCTG 357
```

RESULT 21  
US-10-188-619-14/c  
; Sequence 14, Application US/10186619  
; Publication No. US2003016294A1  
; GENERAL INFORMATION:  
; APPLICANT: Christophe P.G. Gerald, et al.  
; TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2  
; TITLE OF INVENTION: SPECIFIC COMPOUNDS  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/10/188,619  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/407,367  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 44742-AA-PCT-US/JPW  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-10-188-619-14

Query Match 83.2%; Score 15.8; DB 13; Length 45;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGTCTTG 19  
DB 43 CTGGGCATCAGTGTCTTG 25

RESULT 22  
US-09-747-377-447/c  
; Sequence 447, Application US/09747377  
; Publication No. US2003002255A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR BREAST CANCER  
; FILE REFERENCE: A-69959/RMS/DCF  
; CURRENT APPLICATION NUMBER: US/09/747,377  
; CURRENT FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 493  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 447  
; LENGTH: 145  
; TYPE: DNA  
; ORGANISM: Mus sp.  
US-09-747-377-447

Query Match 83.2%; Score 15.8; DB 11; Length 145;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGTCTTG 19  
DB 49 CTGGGCATCAGTGTCTTG 31

RESULT 23  
US-10-105-613-447/c  
; Sequence 447, Application US/10105613  
; Publication No. US20030099963A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR BREAST CANCER  
; FILE REFERENCE: A-69959/RMS/DCF  
; CURRENT APPLICATION NUMBER: US/10/105,613  
; CURRENT FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: US/09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 493  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 447  
; LENGTH: 145  
; TYPE: DNA  
; ORGANISM: Mus sp.  
US-10-105-613-447

Query Match 83.2%; Score 15.8; DB 15; Length 145;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGTCTTG 19  
DB 49 CTGGGCATCAGTGTCTTG 31

RESULT 24  
US-09-920-300A-583  
; Sequence 583, Application US/09920300A  
; Patent No. US20020136728A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Xu, Jianshun  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547  
; CURRENT APPLICATION NUMBER: US/09/920,300A  
; CURRENT FILING DATE: 2001-07-31  
; NUMBER OF SEQ ID NOS: 1789  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 583  
; LENGTH: 422  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-920-300A-583

Query Match 83.2%; Score 15.8; DB 10; Length 422;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGTCTTG 19  
DB 237 CTGGGCATCAGTGTCTTG 255

RESULT 25  
US-10-099-926-583  
; Sequence 583, Application US/10099926  
; Publication No. US2003016064A1

```
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 583
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-099-926-583

Query Match      83.2% Score 15.8; DB 13; Length 422;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTGGGCCATCAGTGTCTG 19
    |||||
Db 237 CTGGGCCATGACTGCTG 255

RESULT 26
US-10-033-528-583
; Sequence 583, Application US/10033528
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 583
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-583

Query Match      83.2% Score 15.8; DB 14; Length 422;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTGGGCCATCAGTGTCTG 19
    |||||
Db 237 CTGGGCCATGACTGCTG 255

RESULT 27
US-10-027-632-281562/c
; Sequence 281562, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; NUMBER OF SEQ ID NOS: 2002-04-30
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281562
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281562

Query Match      83.2% Score 15.8; DB 13; Length 492;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTGGGCCATCAGTGTCTG 19
    |||||
Db 382 CTGGGCCATGACTGCTG 364

RESULT 28
US-10-027-632-281563/c
; Sequence 281563, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; NUMBER OF SEQ ID NOS: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281563
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281563

Query Match      83.2% Score 15.8; DB 13; Length 492;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTGGGCCATCAGTGTCTG 19
    |||||
Db 382 CTGGGCCATGACTGCTG 364

RESULT 29
US-10-027-632-281562/c
; Sequence 281562, Application US/10027632
```

```
GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 281562
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-281562

Query Match
Best Local Similarity 83.2%; Score 15.8; DB 14; Length 492;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGGCGCATGCTGCTG 19
Db 382 CTGGGCGCATGCTGCTG 364

RESULT 30
US-10-027-632-281563/C
; Sequence 281563, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 281563
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-281563

Query Match
Best Local Similarity 83.2%; Score 15.8; DB 14; Length 492;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 CTGGGCGCATGCTGCTG 19
Db 382 CTGGGCGCATGCTGCTG 364

RESULT 31
US-10-027-632-81346
; Sequence 81346, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 81346
; LENGTH: 652
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(652)
; OTHER INFORMATION: n = A,T,C or G
; US-10-027-632-81346

Query Match
Best Local Similarity 83.2%; Score 15.8; DB 13; Length 652;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGGCGCATGCTGCTG 19
Db 339 CAGGGCGCATGAGGCTG 357

RESULT 32
US-10-027-632-81347
; Sequence 81347, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
```

PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 81347  
LENGTH: 652  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(652)  
OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-81347

Query Match 83.2% Score 15.8; DB 13; Length 652;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTGGGCGCATGCTCTG 19  
Db 339 CAGGGCCATCAGGGCTCTG 357

RESULT 33  
US-10-027-632-109771

Sequence 109771, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 109771  
LENGTH: 652  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(652)  
OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-109771

Query Match 83.2% Score 15.8; DB 13; Length 652;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTGGGCGCATGCTCTG 19  
Db 339 CAGGGCCATCAGGGCTCTG 357

RESULT 34  
US-10-027-632-109772

Sequence 109772, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 109772  
LENGTH: 652  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(652)  
OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-109772

Query Match 83.2% Score 15.8; DB 13; Length 652;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTGGGCGCATGCTCTG 19  
Db 339 CAGGGCCATCAGGGCTCTG 357

RESULT 35  
US-10-027-632-301812

Sequence 301812, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 301812  
LENGTH: 652

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(652)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-301812

Query Match      83.2%; Score 15.8; DB 13; Length 652;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 CTGGGCCATCAGTGTCTG 19
      |||||
Db      339 CAGGGCCATCAGGCTCTG 357

RESULT 36
US-10-027-632-301813
; Sequence 301813, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 301813
; LENGTH: 652
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(652)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-301813

Query Match      83.2%; Score 15.8; DB 13; Length 652;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 CTGGGCCATCAGTGTCTG 19
      |||||
Db      339 CAGGGCCATCAGGCTCTG 357

RESULT 37
US-10-027-632-81346
; Sequence 81346, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81346
; LENGTH: 652
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(652)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-81346

Query Match      83.2%; Score 15.8; DB 14; Length 652;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81346
; LENGTH: 652
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(652)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-81347

Query Match      83.2%; Score 15.8; DB 14; Length 652;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 CTGGGCCATCAGTGTCTG 19
      |||||
Db      339 CAGGGCCATCAGGCTCTG 357

RESULT 38
US-10-027-632-81347
; Sequence 81347, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81347
; LENGTH: 652
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(652)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-81347

Query Match      83.2%; Score 15.8; DB 14; Length 652;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```



Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CTGGGCGCATCAGTCTCTG 19  
| | | | | | | | | | | | | | | | | | | | |  
Db 339 CAGGGCGCATCAGGCTCTG 357

RESULT 39  
US-10-027-632-109771  
; Sequence 109771, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 109771  
; LENGTH: 652  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(652)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-109771

Query Match 83.2%; Score 15.8; DB 14; Length 652;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGGCGCATCAGTCTCTG 19  
| | | | | | | | | | | | | | | | | | | | |  
Db 339 CAGGGCGCATCAGGCTCTG 357

RESULT 40  
US-10-027-632-109772  
; Sequence 109772, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 109772  
; LENGTH: 652  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(652)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-109772

Query Match 83.2%; Score 15.8; DB 14; Length 652;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CTGGGCGCATCAGTCTCTG 19  
| | | | | | | | | | | | | | | | | | | | |  
Db 339 CAGGGCGCATCAGGCTCTG 357

Search completed: January 13, 2004, 21:03:12  
Job time : 127.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 13, 2004, 15:53:18 ; Search time 36.5 Seconds  
(without alignments)  
229.761 Million cell updates/sec

Title: US-09-719-737-18

Sequence: 1 ctggcgccatcagtcctcgc 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	100.0	1065	3	US-08-847-296B-2
C 2	19	100.0	1071	4	US-08-567-882-6
C 3	19	100.0	1116	4	US-08-720-565-5
C 4	19	100.0	1193	4	US-08-720-565-3
C 5	19	100.0	1201	4	US-09-016-434-1085
C 6	19	100.0	1689	4	US-08-720-565-1
C 7	19	100.0	1915	3	US-08-575-967A-3
C 8	15.8	83.2	45	2	US-08-687-355A-13
C 9	15.8	83.2	45	4	US-09-407-367-14
C 10	15.8	83.2	1200	4	US-09-016-434-1062
C 11	15.8	83.2	1280	1	US-08-192-288-1
C 12	15.8	83.2	1280	2	US-08-687-355A-1
C 13	15.8	83.2	1280	4	US-09-407-367-1
C 14	15.8	83.2	1786	3	US-09-015-296-2
C 15	15.8	83.2	1786	4	US-09-593-722-2
C 16	15.8	83.2	2144	4	US-08-876-798A-1
C 17	15.8	83.2	2144	4	US-08-876-798A-3
C 18	15.8	83.2	32155	4	US-08-311-731A-1
C 19	15.4	81.1	4403765	3	US-09-103-840A-2
C 20	15.4	81.1	4411529	3	US-09-103-840A-1
C 21	15	78.9	2931	3	US-09-168-629-14
C 22	15	78.9	3966	3	US-09-215-131-1
C 23	15	78.9	3966	3	US-09-222-734-1
C 24	14.8	77.9	45	2	US-08-687-355A-13
C 25	14.8	77.9	45	4	US-09-407-367-13
C 26	14.8	77.9	669	4	US-09-465-901-15
C 27	14.8	77.9	669	4	US-09-465-901-35

C 28	14.8	77.9	1500	4	US-09-685-462-1	Sequence 1, Appl1
C 29	14.8	77.9	1744	4	US-09-685-462-9	Sequence 9, Appl1
C 30	14.8	77.9	2239	4	US-09-581-831-1	Sequence 1, Appl1
C 31	14.8	77.9	2529	4	US-09-620-312D-688	Sequence 688, App
C 32	14.8	77.9	36412	4	US-08-311-731A-132	Sequence 132, App
C 33	14.4	75.8	20	1	US-08-742-023-15	Sequence 15, Appl1
C 34	14.4	75.8	20	3	US-08-968-505-15	Sequence 15, Appl1
C 35	14.4	75.8	264	4	US-09-016-434-481	Sequence 481, App
C 36	14.4	75.8	321	4	US-09-404-879A-327	Sequence 327, App
C 37	14.4	75.8	1071	4	US-09-620-312D-769	Sequence 769, App
C 38	14.2	74.7	20	4	US-09-198-452A-6823	Sequence 6823, Ap
C 39	14.2	74.7	132	3	US-09-461-697-114	Sequence 114, App
C 40	14.2	74.7	141	3	US-09-461-697-112	Sequence 112, App
C 41	14.2	74.7	216	3	US-09-461-697-104	Sequence 104, App
C 42	14.2	74.7	336	3	US-09-461-697-100	Sequence 100, App
C 43	14.2	74.7	510	3	US-09-461-697-96	Sequence 96, Appl1
C 44	14.2	74.7	600	3	US-09-461-697-92	Sequence 92, Appl1
C 45	14.2	74.7	607	3	US-08-889-502-37	Sequence 37, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-847-296B-2/c  
; Sequence 2, Application US/08847296B  
; Patent No. 6271347  
; GENERAL INFORMATION:  
; APPLICANT: DAUGHERTY, BRUCE L.  
; APPLICANT: DEMARTINO, JULIE A.  
; APPLICANT: SICILIANO, SALVATORE J.  
; APPLICANT: SPRINGER, MARTIN J.  
; TITLE OF INVENTION: EOSINOPHIL EOTAXIN RECEPTOR  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merck & Co., Inc.  
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065-0900  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/847,296B  
; FILING DATE: 24-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/016,158  
; FILING DATE: 26-APR-1996  
; APPLICATION NUMBER: 60/017,113  
; FILING DATE: 26-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Thies, J. Eric  
; REGISTRATION NUMBER: 35,382  
; REFERENCE/DOCKET NUMBER: 19634Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 908-594-3904  
; TELEFAX: 908-594-4720  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1065 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; ;  
; MOLECULE TYPE: CDNA  
; ;  
; US-08-847-296B-2  
Query Match 100.0%; Score 19; DB 3; Length 1065;

Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGGCGCATCAGTCTCTG 19  
Db 105 CTGGGCGCATCAGTCTCTG 87

## RESULT 2

US-08-567-882-6/c  
Sequence 6, Application US/08567882  
Patent No. 6512103  
GENERAL INFORMATION:  
APPLICANT: Dairaghi, Daniel J.  
APPLICANT: Hara, Takahiko  
APPLICANT: Miyajima, Atsushi  
APPLICANT: Schall, Thomas J.  
APPLICANT: Wang, Wei  
APPLICANT: Yoshimura, Akihiko  
TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/567,882  
FILING DATE: 08-DEC-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0506  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1071 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1071  
US-08-567-882-6

Query Match 100.0%; Score 19; DB 4; Length 1071;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGGCGCATCAGTCTCTG 19  
Db 105 CTGGGCGCATCAGTCTCTG 87

## RESULT 3

US-08-720-565-5/c  
Sequence 5, Application US/08720565  
Patent No. 6537764  
GENERAL INFORMATION:  
APPLICANT: Gerard, Craig J.  
APPLICANT: Gerard, No. 6537764ma P.  
APPLICANT: Mackay, Charles R.

APPLICANT: Ponath, Paul D.  
APPLICANT: Post, Theodore W.  
APPLICANT: Qin, Shixin  
TITLE OF INVENTION: G PROTEIN-COUPLED RECEPTOR GENE CCR3 AND  
TITLE OF INVENTION: ANTAGONISTS THEREOF  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,565  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/00608  
FILING DATE: 19-JAN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375,199  
FILING DATE: 19-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: LKS94-05A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1116 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-720-565-5

Query Match 100.0%; Score 19; DB 4; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 0.88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGGCGCATCAGTCTCTG 19  
Db 119 CTGGGCGCATCAGTCTCTG 101

## RESULT 4

US-08-720-565-3/c  
Sequence 3, Application US/08720565  
Patent No. 6537764  
GENERAL INFORMATION:  
APPLICANT: Gerard, Craig J.  
APPLICANT: Gerard, No. 6537764ma P.  
APPLICANT: Mackay, Charles R.  
APPLICANT: Ponath, Paul D.  
APPLICANT: Post, Theodore W.  
APPLICANT: Qin, Shixin  
TITLE OF INVENTION: G PROTEIN-COUPLED RECEPTOR GENE CCR3 AND  
TITLE OF INVENTION: ANTAGONISTS THEREOF  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Releasee #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,565  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/00608  
FILING DATE: 19-JAN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375,199  
FILING DATE: 19-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: LKS94-05A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1193 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 92..1156  
US-08-720-565-3

Query Match 100.0%; Score 19; DB 4; Length 1193;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCGATCAGTCTCTG 19  
DB 196 CTGGGCGATCAGTCTCTG 178

RESULT 5  
US-09-016-434-1085/C  
Sequence 1085, Application US/09016434  
Patent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Selthamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1085:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1201 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g1199579  
US-09-016-434-1085

Query Match 100.0%; Score 19; DB 4; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCGATCAGTCTCTG 19  
DB 136 CTGGGCGATCAGTCTCTG 118

RESULT 6  
US-08-720-565-1/C  
Sequence 1, Application US/08720565  
Patent No. 6537764  
GENERAL INFORMATION:  
APPLICANT: Gerard, Craig J.  
APPLICANT: Mackay, Charles R.  
APPLICANT: Ponath, Paul D.  
APPLICANT: Post, Theodore W.  
APPLICANT: Qin, Shixin  
TITLE OF INVENTION: G PROTEIN-COUPLED RECEPTOR GENE CCR3 AND  
TITLE OF INVENTION: ANTAGONISTS THEREOF  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millita Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Releasee #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,565  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/00608  
FILING DATE: 19-JAN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375,199  
FILING DATE: 19-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: LKS94-05A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1689 base pairs

TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-720-565-1

Query Match 100.0%; Score 19; DB 4; Length 1689;  
Best Local Similarity 100.0%; Pred. No. 0.93;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTCTG 19  
Db 285 CTGGGCCATCAGTGTCTCTG 267

RESULT 7  
US-08-575-967A-3/c  
Sequence 3, Application US/08575967A  
Patent No. 6265184

GENERAL INFORMATION:  
APPLICANT: Gray et al.  
TITLE OF INVENTION: Chemokine Receptor Materials and Methods  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PC-DOS/MS-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/575,967A  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: No. 6265184and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32918  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-485-1900  
TELEFAX: 206-485-1662

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1915 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 362..1426  
NAME/KEY: misc\_feature  
OTHER INFORMATION: /= "88-2B polynucleotide and amino acid  
US-08-575-967A-3

Query Match 100.0%; Score 19; DB 3; Length 1915;  
Best Local Similarity 100.0%; Pred. No. 0.94;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTCTG 19  
Db 466 CTGGGCCATCAGTGTCTCTG 448

RESULT 8  
US-08-687-355A-14/c  
Sequence 14, Application US/08687355A

Patent No. 5989834  
GENERAL INFORMATION:  
APPLICANT: Synaptic Pharmaceutical Corporation  
TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE  
TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,355A  
FILING DATE: No. 5989834ember 26, 1996  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPW/MAT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-378-0400  
TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-687-355A-14

Query Match 83.2%; Score 15.8; DB 2; Length 45;  
Best Local Similarity 89.5%; Pred. No. 25;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTCTG 19  
Db 43 CTGGGCCATCAGTGTCTCTG 25

RESULT 9  
US-09-407-367-14/c  
Sequence 14, Application US/09407367  
Patent No. 6420532

GENERAL INFORMATION:  
APPLICANT: Christophe P.G. Gerald, et al.  
TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2 SPECIFIC COMPO  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/407,367  
FILING DATE:  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 44742-AA-PCT-US/JPM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-407-367-14

Query Match 83.2%; Score 15.8; DB 4; Length 45;  
Best Local Similarity 89.5%; Pred. No. 25;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGTCTG 19  
DB 43 CTGGGCATCAGTGTCTG 25

RESULT 10  
US-09-016-434-1062  
Sequence 1062, Application US/09016434  
Patent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016.434  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1062:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1200 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g1063633

US-09-016-434-1062

Query Match 83.2%; Score 15.8; DB 4; Length 1200;  
Best Local Similarity 89.5%; Pred. No. 39;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGTCTG 19  
DB 536 CTGGGCATCAGTGTCTG 554

RESULT 11  
US-08-192-288-1  
Sequence 1, Application US/08192288  
Patent No. 5545549  
GENERAL INFORMATION:  
APPLICANT: Gerald, Christophe  
APPLICANT: Walker, Mary  
APPLICANT: Branchek, Theresa  
APPLICANT: Weinsbank, Richard L.  
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE  
TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTOR AND USES  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/192,288  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 44742/JPM/TEP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1280 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 43..1185  
US-08-192-288-1

Query Match 83.2%; Score 15.8; DB 1; Length 1280;  
Best Local Similarity 89.5%; Pred. No. 39;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTGTCTG 19  
DB 558 CTGGGCATCAGTGTCTG 576

RESULT 12  
US-08-687-355A-1  
Sequence 1, Application US/08687355A  
Patent No. 5989834

GENERAL INFORMATION:  
APPLICANT: Synaptic Pharmaceutical Corporation  
TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE  
TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,355A  
FILING DATE: No. 5989834ember 26, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPM/MAT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1280 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 43..1185  
US-08-687-355A-1

Query Match 83.2%; Score 15.8; DB 2; Length 1280;  
Best Local Similarity 89.5%; Pred. No. 39;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGGCATCAGTCTCTG 19  
DB 558 CTGGGGCATCAGTCTCTG 576

RESULT 13  
US-09-407-367-1  
Sequence 1, Application US/09407367  
Patent No. 6420532  
GENERAL INFORMATION:  
APPLICANT: Christophe P.G. Gerald, et al.  
TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2 SPECIFIC COMPO  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/407,367  
FILING DATE:  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 44742-AA-PCT-US/JPM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1280 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 43..1185  
US-09-407-367-1

Query Match 83.2%; Score 15.8; DB 4; Length 1280;  
Best Local Similarity 89.5%; Pred. No. 39;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGGCATCAGTCTCTG 19  
DB 558 CTGGGGCATCAGTCTCTG 576

RESULT 14  
US-09-015-296-2/c  
Sequence 2, Application US/09015296  
Patent No. 6103471  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: HUMAN BETA-ALANINE-PYRUVATE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Potter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/015,296  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0467 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1786 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:



LIBRARY: LUNGTUT12  
CLONE: 3128715  
US-09-015-296-2

Query Match 83.2%; Score 15.8; DB 3; Length 1786;  
Best Local Similarity 89.5%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTG 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1270 CAGGGCCATCAGTGTCTG 1252

## RESULT 15

US-09-593-722-2/c  
Sequence 2, Application US/09593722  
Patent No. 6416755  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
Lal, Preeti  
Corley, Neil C.  
TITLE OF INVENTION: HUMAN BETA-ALANINE-PYRUVATE  
AMINOTRANSFERASE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/593,722  
FILING DATE: 13-Jun-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/015,296  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0467 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1786 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LUNGTUT12  
CLONE: 3128715  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-593-722-2

Query Match 83.2%; Score 15.8; DB 4; Length 1786;  
Best Local Similarity 89.5%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTG 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1270 CAGGGCCATCAGTGTCTG 1252

## RESULT 16

US-08-876-798A-1  
Sequence 1, Application US/08876798A  
Patent No. 6355478  
GENERAL INFORMATION:

APPLICANT: Baez, Melvyn  
Yang, Peiyi  
TITLE OF INVENTION: RHESUS MONKEY NEUROPEPTIDE Y RECEPTOR  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.A.  
ZIP: 46285

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/876,798A  
FILING DATE: 16-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gaylo, Paul J.  
REGISTRATION NUMBER: 36,808  
REFERENCE/DOCKET NUMBER: X-10901  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317)276-0756  
TELEFAX: (317)276-2763

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2144 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 63..1205  
US-08-876-798A-1

Query Match 83.2%; Score 15.8; DB 4; Length 2144;  
Best Local Similarity 89.5%; Pred. No. 42;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTG 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 578 CTGGGCCATCAGTGTCTG 596

## RESULT 17

US-08-876-798A-3  
Sequence 3, Application US/08876798A  
Patent No. 6355478  
GENERAL INFORMATION:  
APPLICANT: Baez, Melvyn  
Yang, Peiyi  
TITLE OF INVENTION: RHESUS MONKEY NEUROPEPTIDE Y RECEPTOR  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.A.  
ZIP: 46285

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/876,798A  
FILING DATE: 16-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gaylo, Paul J.  
REGISTRATION NUMBER: 36,808  
REFERENCE/DOCKET NUMBER: X-10901  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317)276-0756  
TELEFAX: (317)276-2763  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2144 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: mRNA  
US-08-876-798A-3

Query Match 83.2%; Score 15.8; DB 4; Length 2144;  
Best Local Similarity 68.4%; Pred. No. 42;  
Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTGCTCTG 19  
|||||  
Db 578 CTGGGCAUCAGUGCCUG 596

RESULT 18  
US-08-311-731A-1  
Sequence 1, Application US/08311731A  
Patent No. 6583266  
GENERAL INFORMATION:  
APPLICANT: SMITH, DOUGLAS  
APPLICANT: MAO, JEN-I  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 411  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
STREET: 600 ATLANTIC AVENUE  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,731A  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: GATES, EDWARD R.  
REGISTRATION NUMBER: 31,616  
REFERENCE/DOCKET NUMBER: C0044/7125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/720-3500  
TELEFAX: 617/720-2441  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32155 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:

ORGANISM: MYCOBACTERIUM TUBERCULOSIS  
US-08-311-731A-1

Query Match 81.1%; Score 15.4; DB 4; Length 32155;  
Best Local Similarity 94.1%; Pred. No. 97;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGGGCCATCAGTGCTCT 18  
|||||  
Db 25340 TGGGCCATCGGTGCTCT 25356

RESULT 19  
US-09-103-840A-2/c  
Sequence 2, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
TUBERCULOSIS  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4403765  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
OTHER INFORMATION: CDC 1551  
OTHER INFORMATION: "n" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 81.1%; Score 15.4; DB 3; Length 4403765;  
Best Local Similarity 94.1%; Pred. No. 84;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGGGCCATCAGTGCTCT 18  
|||||  
Db 3293522 TGGGCCATCGGTGCTCT 3293506

RESULT 20  
US-09-103-840A-1/c  
Sequence 1, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
TUBERCULOSIS  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37RV  
US-09-103-840A-1

Query Match 81.1%; Score 15.4; DB 3; Length 4411529;  
Best Local Similarity 94.1%; Pred. No. 84;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGGGCCATCAGTCTCT 18  
Db 3299247 TGGGCCATCAGTCTCT 3299231

RESULT 21  
US-09-168-629-14/C  
; Sequence 14, Application US/09168629  
; Patent No. 6242253  
; GENERAL INFORMATION:  
; APPLICANT: Karin, Michael  
; APPLICANT: Didonato, Joseph A.  
; APPLICANT: Rochwarf, David M.  
; APPLICANT: Hayakawa, Makio  
; APPLICANT: Zandi, Ebrahim  
; TITLE OF INVENTION: IKK Kinase, Subunits Thereof, and Methods of Using Same  
; FILE REFERENCE: P-UD 3295  
; CURRENT APPLICATION NUMBER: US/09/168,629  
; CURRENT FILING DATE: 1998-10-08  
; EARLIER APPLICATION NUMBER: 60/061,470  
; EARLIER FILING DATE: 1997-10-09  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 2931  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (36)..(2306)  
US-09-168-629-14

Query Match 78.9%; Score 15; DB 3; Length 2931;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGGGCCATCAGTGC 15  
Db 2468 CTTGGGCCATCAGTGC 2454

RESULT 22  
US-09-215-131-1/C  
; Sequence 1, Application US/09215131  
; Patent No. 6030834  
; GENERAL INFORMATION:  
; APPLICANT: Chu, Keting  
; APPLICANT: Pot, David  
; TITLE OF INVENTION: IKK Beta Regulates Transcription Factors  
; FILE REFERENCE: 1449,002  
; CURRENT APPLICATION NUMBER: US/09/215,131  
; CURRENT FILING DATE: 1998-12-18  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 3966  
; TYPE: DNA  
; ORGANISM: human  
US-09-215-131-1

Query Match 78.9%; Score 15; DB 3; Length 3966;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGGGCCATCAGTGC 15  
Db 2433 CTTGGGCCATCAGTGC 2419

RESULT 23  
US-09-222-734-1/C  
; Sequence 1, Application US/09222734A

; Patent No. 6077701  
; GENERAL INFORMATION:  
; APPLICANT: Chu, Keting  
; APPLICANT: Pot, David  
; TITLE OF INVENTION: IKK-beta Regulates Transcription Factors  
; FILE REFERENCE: 12441,78080  
; CURRENT APPLICATION NUMBER: US/09/222,734A  
; CURRENT FILING DATE: 1998-12-29  
; EARLIER APPLICATION NUMBER: 09/215,131  
; EARLIER FILING DATE: 1998-12-18  
; EARLIER APPLICATION NUMBER: 60/068,954  
; EARLIER FILING DATE: 1997-12-30  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3966  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-222-734-1

Query Match 78.9%; Score 15; DB 3; Length 3966;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGGGCCATCAGTGC 15  
Db 2433 CTTGGGCCATCAGTGC 2419

RESULT 24  
US-08-687-355A-13  
; Sequence 13, Application US/08687355A  
; Patent No. 5989834  
; GENERAL INFORMATION:  
; APPLICANT: Symplic Pharmaceuticals Corporation  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE  
; TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/687,355A  
; FILING DATE: No. 5989834ember 26, 1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPW/MAT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
US-08-687-355A-13

Query Match 77.9%; Score 14.8; DB 2; Length 45;

Best Local Similarity 88.9%; Pred. No. 80;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTCTCT 18  
DB 28 CTGGGCATCAGTCTCT 45

## RESULT 25

US-09-407-367-13

; Sequence 13, Application US/09407367  
; Patent No. 6420532  
; GENERAL INFORMATION:  
; APPLICANT: Christophe P.G. Gerald, et al.  
; TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2 SPECIFIC COMPO  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/407,367  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 44742-AA-PCT-US/JPM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; US-09-407-367-13

Query Match 77.9%; Score 14.8; DB 4; Length 45;  
Best Local Similarity 88.9%; Pred. No. 80;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGGCATCAGTCTCT 18  
DB 28 CTGGGCATCAGTCTCT 45

## RESULT 26

US-09-465-901-15

; Sequence 15, Application US/09465901  
; Patent No. 6492143  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Randall  
; APPLICANT: Yau, King-Wai  
; TITLE OF INVENTION: Olfactory Receptor Expression Libraries  
; TITLE OF INVENTION: ad Methods of Making and Using Them  
; FILE REFERENCE: 001107.00105  
; CURRENT APPLICATION NUMBER: US/09/465,901  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: 60/112,605

; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15

; LENGTH: 669  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: PCR primer  
; US-09-465-901-15

Query Match 77.9%; Score 14.8; DB 4; Length 669;  
Best Local Similarity 88.9%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGGGCATCAGTCTCTG 19  
DB 256 TGATCATCTGTCTCTG 273

## RESULT 27

US-09-465-901-35

; Sequence 35, Application US/09465901  
; Patent No. 6492143  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Randall  
; APPLICANT: Yau, King-Wai  
; APPLICANT: Krautwurst, Dietmar  
; TITLE OF INVENTION: Olfactory Receptor Expression Libraries  
; TITLE OF INVENTION: ad Methods of Making and Using Them  
; FILE REFERENCE: 001107.00105  
; CURRENT APPLICATION NUMBER: US/09/465,901  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: 60/112,605  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 669  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: PCR primer  
; US-09-465-901-35

Query Match 77.9%; Score 14.8; DB 4; Length 669;  
Best Local Similarity 88.9%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGGGCATCAGTCTCTG 19  
DB 256 TGATCATCTGTCTCTG 273

## RESULT 28

US-09-685-462-1/c

; Sequence 1, Application US/09685462  
; Patent No. 6524833  
; GENERAL INFORMATION:  
; APPLICANT: Zon, Leonard I.  
; APPLICANT: Agarwal, Sachana  
; APPLICANT: Best, Jennifer  
; APPLICANT: Vail, Brenda  
; TITLE OF INVENTION: Two Sterile-20 Kinase-Like Proteins and  
; FILE REFERENCE: 1242.1016-004  
; CURRENT APPLICATION NUMBER: US/09/685,462  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/591,083  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: PCT/US98/26116  
; PRIOR FILING DATE: 1998-12-09  
; PRIOR APPLICATION NUMBER: US 60/069,078

PRIOR FILING DATE: 1997-12-09  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1500  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (20)...(1258)  
US-09-685-462-1

Query Match 77.9%; Score 14.8; DB 4; Length 1500;  
Best Local Similarity 88.9%; Pred. No. 1.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGGGCCATCAGTCTCTG 19  
DB 1294 TGGCCCTTCAGTGCCTG 1277

RESULT 29  
US-09-685-462-9/c  
Sequence 9, Application US/09685462  
Patent No. 6524833  
GENERAL INFORMATION:  
APPLICANT: Zon, Leonard I.  
APPLICANT: Agarwal, Sadhana  
APPLICANT: Best, Jennifer  
APPLICANT: Vail, Brenda  
TITLE OF INVENTION: Two Sterile-20 Kinase-Like Proteins and  
TITLE OF INVENTION: Methods of Use Thereof  
FILE REFERENCE: 1242.1016-004  
CURRENT APPLICATION NUMBER: US/09/685,462  
CURRENT FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/591,083  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: PCT/US98/26116  
PRIOR FILING DATE: 1998-12-09  
PRIOR APPLICATION NUMBER: US 60/069,078  
PRIOR FILING DATE: 1997-12-09  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 1744  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (302)...(1597)  
US-09-685-462-9

Query Match 77.9%; Score 14.8; DB 4; Length 1744;  
Best Local Similarity 88.9%; Pred. No. 1.3e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGGGCCATCAGTCTCTG 19  
DB 1604 TGGCCCTTCAGTGCCTG 1587

RESULT 30  
US-09-581-831-1  
Sequence 1, Application US/09581831  
Patent No. 6448020  
GENERAL INFORMATION:  
APPLICANT: TOFTGARD, RUNE  
APPLICANT: ZAPHIROPOULOS, PETER G.  
APPLICANT: KOGERMAN, PRIT  
APPLICANT: GRIMM, THOMAS  
TITLE OF INVENTION: MOLECULES ASSOCIATED WITH THE HUMAN SUPPRESSOR OF FUSED  
TITLE OF INVENTION: GENE  
FILE REFERENCE: 50695-60568

CURRENT APPLICATION NUMBER: US/09/581,831  
CURRENT FILING DATE: 2000-08-21  
PRIOR APPLICATION NUMBER: PCT/SE98/02383  
PRIOR FILING DATE: 1998-12-18  
PRIOR APPLICATION NUMBER: 9704788-0  
PRIOR FILING DATE: 1997-12-19  
PRIOR APPLICATION NUMBER: 9802293-2  
PRIOR FILING DATE: 1998-06-26  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 2239  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (167)...(1618)  
US-09-581-831-1

Query Match 77.9%; Score 14.8; DB 4; Length 2239;  
Best Local Similarity 88.9%; Pred. No. 1.4e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTCTCT 18  
DB 98 CTGGCCGTCAGTCTCT 115

RESULT 31  
US-09-620-312D-688/c  
Sequence 688, Application US/09620312D  
Patent No. 6569662  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Weinman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yundong  
APPLICANT: Wang, Duntui  
APPLICANT: Wang, Zhilwei  
APPLICANT: John Tillinghaast  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pc\_pl\_genes Version 1.0  
SEQ ID NO 688  
LENGTH: 2529  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (169)...(1134)  
US-09-620-312D-688

Query Match 77.9%; Score 14.8; DB 4; Length 2529;  
Best Local Similarity 88.9%; Pred. No. 1.4e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CTGGGCGCATGCTCT 18  
Db 1433 CTGGGCGCATGCTCT 1416

RESULT 32  
US-08-311-731A-132/c  
; Sequence 132, Application US/08311731A  
; Patent No. 6583266  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, DOUGLAS  
; APPLICANT: MAO, JEN-I  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 411  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
; STREET: 600 ATLANTIC AVENUE  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,731A  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GATES, EDWARD R.  
; REGISTRATION NUMBER: 31,616  
; REFERENCE/DOCKET NUMBER: C0044/7125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/720-3500  
; TELEFAX: 617/720-2441  
; INFORMATION FOR SEQ ID NO: 132:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 36412 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; MOLECULE TYPE: circular  
; TOPOLOGY: circular  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: MYCOBACTERIUM LEPRAE  
; US-08-311-731A-132

Query Match 77.9%; Score 14.8; DB 4; Length 36412;  
Best Local Similarity 88.9%; Pred. No. 2e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TGGGCGCATGCTCTG 19  
Db 11540 TGGGCGCATGCTGCGCTG 11523

RESULT 33  
US-08-742-023-15  
; Sequence 15, Application US/08742023  
; Patent No. 5800997  
; GENERAL INFORMATION:  
; APPLICANT: Beck, James J.  
; TITLE OF INVENTION: Detection of Maize Fungal Pathogens  
; TITLE OF INVENTION: Using the Polymerase Chain Reaction  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation

STREET: 520 White Plains Road, P.O. Box 2005  
; CITY: Tarrytown  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10591  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/742,023  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919) 541-8587  
; TELEFAX: (919) 541-8689  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "primer JBS86"  
; US-08-742-023-15

Query Match 75.8%; Score 14.4; DB 1; Length 20;  
Best Local Similarity 93.8%; Pred. No. 1.2e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 GCGCATGCTCTG 19  
Db 2 GCGCATGCTCTG 17

RESULT 34  
US-08-968-505-15  
; Sequence 15, Application US/08968505  
; Patent No. 6071698  
; GENERAL INFORMATION:  
; APPLICANT: Beck, James J.  
; TITLE OF INVENTION: Detection of Maize Fungal Pathogens  
; TITLE OF INVENTION: Using the Polymerase Chain Reaction  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 520 White Plains Road, P.O. Box 2005  
; CITY: Tarrytown  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10591  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/968,505  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/742,023  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919) 541-8587  
; TELEFAX: (919) 541-8689

INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "primer JB586"  
US-08-968-505-15

Query Match 75.8%; Score 14.4; DB 3; Length 20;  
Best Local Similarity 93.8%; Pred. No. 1.2e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGCCATCAGTCTCTG 19  
DB 2 GGCCATCAGTCTCTG 17

## RESULT 35

US-09-016-434-481  
Sequence 481, Application US/09016434  
Patent No. 6500938

GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Sellhammer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 481:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: UTRSN0202  
CLONE: 2260261  
US-09-016-434-481

Query Match 75.8%; Score 14.4; DB 4; Length 264;  
Best Local Similarity 93.8%; Pred. No. 1.6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGCCATCAGTCTCTG 19  
DB 2 GGCCATCAGTCTCTG 17

DB 38 GGCCATCAGTCCCTG 53

## RESULT 36

US-09-404-879A-327/c  
Sequence 327, Application US/09404879A  
Patent No. 6468546  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462C2  
CURRENT APPLICATION NUMBER: US/09/404,879A  
CURRENT FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 393  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 327  
LENGTH: 321  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-404-879A-327

Query Match 75.8%; Score 14.4; DB 4; Length 321;  
Best Local Similarity 93.8%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGCCATCAGTCTCTG 19  
DB 210 GGCCATCAGTCCCTG 195

## RESULT 37

US-09-620-312D-769/c  
Sequence 769, Application US/09620312D  
Patent No. 6569662

GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yundong  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhilwei  
APPLICANT: John Tillinghaast  
APPLICANT: Drmanac, Radote T.  
TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pc\_fl\_genes Version 1.0  
SEQ ID NO 769  
LENGTH: 1071  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (120)...(890)

US-09-620-312D-769

Query Match 75.8%; Score 14.4; DB 4; Length 1071;  
Best Local Similarity 93.8%; Pred. No. 2e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGCCATCAGTGTCTG 19  
|||||  
Db 498 GGCCATCAGGCTCTG 483

RESULT 38

US-09-198-452A-6823  
; Sequence 6823, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffiths, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 6823  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-6823

Query Match 74.7%; Score 14.2; DB 4; Length 20;  
Best Local Similarity 84.2%; Pred. No. 1.5e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTGTCTG 19  
|||||  
Db 1 CTGGACCATAGTACTCTG 19

RESULT 39

US-09-461-697-114/c  
; Sequence 114, Application US/09461697  
; Patent No. 6277974  
; GENERAL INFORMATION:  
; APPLICANT: COGENT NEUROSCIENCE, Inc.  
; APPLICANT: Lo, Donald C.  
; APPLICANT: Barney, Shawn  
; APPLICANT: Thomas, Mary Beth  
; APPLICANT: Portbury, Stuart D.  
; APPLICANT: Puranam, Kasturi  
; APPLICANT: Katz, Lawrence C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING  
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING  
; TITLE OF INVENTION: CELL DEATH  
; FILE REFERENCE: 10001-005-999  
; CURRENT APPLICATION NUMBER: US/09/461,697  
; CURRENT FILING DATE: 1999-12-14  
; NUMBER OF SEQ ID NOS: 466  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 114  
; LENGTH: 132  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-461-697-114

Query Match 74.7%; Score 14.2; DB 3; Length 132;  
Best Local Similarity 84.2%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTGTCTG 19  
|||||  
Db 40 CTGAGTCATCAGGCTCTG 22

RESULT 40

US-09-461-697-112/c  
; Sequence 112, Application US/09461697  
; Patent No. 6277974  
; GENERAL INFORMATION:  
; APPLICANT: COGENT NEUROSCIENCE, Inc.  
; APPLICANT: Lo, Donald C.  
; APPLICANT: Barney, Shawn  
; APPLICANT: Thomas, Mary Beth  
; APPLICANT: Portbury, Stuart D.  
; APPLICANT: Puranam, Kasturi  
; APPLICANT: Katz, Lawrence C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING  
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING  
; FILE REFERENCE: 10001-005-999  
; CURRENT APPLICATION NUMBER: US/09/461,697  
; CURRENT FILING DATE: 1999-12-14  
; NUMBER OF SEQ ID NOS: 466  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 112  
; LENGTH: 141  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-461-697-112

Query Match 74.7%; Score 14.2; DB 3; Length 141;  
Best Local Similarity 84.2%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCCATCAGTGTCTG 19  
|||||  
Db 49 CTGAGTCATCAGGCTCTG 31

Search completed: January 13, 2004, 17:58:45  
Job time : 47.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: January 13, 2004, 15:53:18 ; Search time 123.5 Seconds  
(without alignments)  
415.298 Million cell updates/sec

Title: US-09-719-737-18

Perfect score: 1 ctcggccatcagctgctg 19

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N\_Geneseq\_19Jun03.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	AAZ57602	Antisense oligonuc
2	19	100.0	19	ABX12695	Human CCR3 recepto
3	19	100.0	1068	AAZ21266	Human low adenosin
4	19	100.0	1068	AAA35144	Human adenosine re
5	19	100.0	1068	AB197977	Non-endogenous hum
6	19	100.0	1068	ABA94340	Human CC chemokine
7	19	100.0	1068	AAZ25222	Human chemokine (C
8	19	100.0	1071	AAZ79096	Human CCR3 chemok

C	9	19	100.0	1116	17	AAZ31336	CC-chemokine recep
C	10	19	100.0	1116	19	AAV07404	Human C-C chemokin
C	11	19	100.0	1193	17	AAZ31335	CC-chemokine recep
C	12	19	100.0	1193	19	AAV07403	Human C-C chemokin
C	13	19	100.0	1201	21	AAZ21267	Human low adenosin
C	14	19	100.0	1201	21	AAA35145	Human adenosine re
C	15	19	100.0	1201	24	ABR84282	Human CDNA differe
C	16	19	100.0	1201	25	ACB56487	Human signaling p
C	17	19	100.0	1201	25	ABZ42635	Human C-C chemokin
C	18	19	100.0	1689	17	AAZ31334	CC-chemokine recep
C	19	19	100.0	1689	18	AAZ58783	Human C-C chemokin
C	20	19	100.0	1689	19	AAV07402	Human C-C chemokin
C	21	19	100.0	1689	21	AAZ21268	Human low adenosin
C	22	19	100.0	1689	21	AAA35146	Human adenosine re
C	23	19	100.0	1689	24	ABL40462	Human C-C chemokin
C	24	19	100.0	1689	25	ABX13645	Human CDNA encodin
C	25	19	100.0	1717	24	ABZ67066	Thyroid cancer rel
C	26	19	100.0	1717	24	AAZ25221	Human chemokine (C
C	27	19	100.0	1717	24	AAZ25245	Human chemokine (C
C	28	19	100.0	1915	18	AAZ85162	Human chemokine re
C	29	19	100.0	3426	24	ABT04010	Human ovary specif
C	30	19	100.0	3958	21	AAZ21269	Human low adenosin
C	31	19	100.0	3958	21	AAA35147	Human adenosine re
C	32	19	100.0	5089	18	AAZ3601	Human eosinophil e
C	33	19	100.0	5791	25	ABZ68879	Nucleotide sequenc
C	34	17.4	91.6	549	22	ABA31445	Human foetal liver
C	35	17.4	91.6	549	22	ABA31445	Probe #9611 for ge
C	36	17.4	91.6	549	22	AAK12471	Human brain expres
C	37	17.4	91.6	549	22	AAK38179	Human bone marrow
C	38	17.4	91.6	549	22	AAI18973	Probe #8906 for ge
C	39	17.4	91.6	549	22	AAI44108	Probe #12794 used
C	40	17.4	91.6	549	23	ABZ37802	Human liver single
C	41	17.4	91.6	549	24	ABZ12213	Human genome-deriv
C	42	17.4	91.6	1565	22	AAZ86115	APEX-2 CDNA. Mus
C	43	17.4	91.6	3728	22	AAZ89968	Human bone marrow
C	44	16.4	86.3	625	22	AAZ31539	Human olfactory re
C	45	15.8	83.2	45	16	AAQ95044	Human hippocampal

ALIGNMENTS

RESULT 1  
AAZ57602  
ID AAZ57602 standard; DNA; 19 BP.

AC AAZ57602;  
XX 28-MAR-2000 (first entry)  
DE Antisense oligonucleotide CCR3AS to inhibit CCR3 receptor expression.  
XX  
XX Antisense oligonucleotide; CCR3 receptor; chemokine receptor; asthma;  
KW allergy; cancer; receptor expression inhibitor; hyperesinophilia;  
XX inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO966037-A2.  
XX  
XX 23-DEC-1999.  
XX  
XX 17-JUN-1999; 99WO-CN00572.  
XX  
XX 17-JUN-1998; 98CA-2235420.  
XX  
XX  
XX (REEX-) RECH EXPERTISES & DEV MEDICAUX PAREN2 IN.  
XX Renz P;  
XX WPI; 2000-097743/08.  
XX  
XX Antisense oligonucleotides directed to CCR3, interleukin or granulocyte

PT macrophage colony stimulating factor receptors, used for treating or  
PT preventing asthma, allergies, hyper eosinophilia, inflammation or cancer  
PS  
XX  
PS Claim 5; Page 32; 72pp; English.  
XX  
CC This is an antisense oligonucleotide directed against the CCR3 receptor.  
CC The antisense oligonucleotide inhibits CCR3 receptor expression. The CCR3  
CC receptor is important in the recruitment of eosinophils into the sites of  
CC allergic or asthmatic inflammation. The chemokines RANTES, MCP-4 and  
CC RANTES mediate most of their effects through the CCR3 receptor. The  
CC invention relates to antisense oligonucleotides directed against a  
CC nucleic acid sequence encoding either a CCR3 receptor, a common subunit  
CC of interleukin-4 (IL-4) and interleukin-13 (IL-13) receptors, or a common  
CC subunit of IL-3, IL-5 and GM-CSF receptors. The antisense  
CC oligonucleotides can be used in the treatment or prevention of asthma,  
CC allergy, hyper eosinophilia, general inflammation or cancer.  
XX  
SQ Sequence 19 BP; 2 A; 6 C; 6 G; 5 T; 0 other;  
Query Match 100.0%; Score 19; DB 21; Length 19;  
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Gaps 0;  
Matches 19; Conservative 0; Indels 0; Gaps 0;  
QY 1 CTGGGCCATCAGTCTCTG 19  
1 CTGGGCCATCAGTCTCTG 19  
DB 1 CTGGGCCATCAGTCTCTG 19  
RESULT 2  
ABX12695  
ID ABX12695 standard; DNA; 19 BP.  
XX  
AC ABX12695;  
XX  
DT 10-MAY-2003 (first entry)  
XX  
DE Human CCR3 receptor DNA, antisense oligonucleotide #1.  
XX  
KW Human; inflammation; 2',6'-diaminopurine; DAP; antisense therapy;  
KW DAP-modified oligonucleotide; pulmonary disease; respiratory disease;  
KW neurological disease; cardiovascular disease; rheumatological disease;  
KW urticarial disease; cutaneous disease; ophthalmological disease;  
KW urinary system disease; pathogen infection; genetic disease; cancer;  
KW airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;  
KW cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;  
KW eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;  
KW hyper eosinophilia; carotid; ophthalmological; immunosuppressive;  
KW antiaesthetic; antiallergic; antiinflammatory; immunosuppressive;  
KW atopic disease; neoplastic cell proliferation; antisense;  
KW CCR3 receptor; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200304511-A2.  
XX  
PD 16-JAN-2003.  
XX  
PF 08-JUL-2002; 2002WO-CA01046.  
XX  
PR 06-JUL-2001; 2001US-303071P.  
XX  
PA (TOP1-) TOP1GEN PHARM INC.  
XX  
PI Renzi P, Allam M, Alakheridi Z;  
XX  
XX WPI; 2003-247944/24.  
PT Increasing in vivo efficacy of a nucleic acid molecule that is  
PT administered to a mammal for inhibiting inflammation in mammals,  
PT involves incorporating into the nucleic acid molecule at least one  
PT nucleotide substitute -  
XX

PS Claim 28; Page 11; 63pp; English.  
XX  
XX The present invention relates to a method for increasing the in vivo  
CC efficacy of oligonucleotides and inhibiting inflammation. The  
CC oligonucleotides comprise at least one nucleotide substitute of  
CC 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide  
CC substitutions are useful for increasing in vivo efficacy of a nucleic  
CC acid molecule that is administered to a mammal. The DAP-modified  
CC oligonucleotides are useful in antisense therapy for treating and/or  
CC preventing pulmonary/respiratory diseases, neurological diseases,  
CC cardiovascular diseases, rheumatological diseases, digestive diseases,  
CC cutaneous diseases, ophthalmological diseases, urinary system diseases,  
CC pathogen infections, genetic diseases, general inflammation and  
CC cancers. The respiratory system disease is a sickness associated with  
CC an inflammation of the lungs, the airways and/or the nose. The  
CC respiratory system disease is selected from pulmonary fibrosis, adult  
CC respiratory distress syndrome, cystic fibrosis, chronic obstructive  
CC lung disease, chronic bronchitis, eosinophilic bronchitis, asthma,  
CC allergy, allergic rhinitis, sinusitis and hyper eosinophilia. The  
CC DAP-modified oligonucleotides are more stable in the body, more  
CC effective, and less toxic than standard antisense oligonucleotides.  
CC DAP or its analogues are more effective than other substitutes of  
CC adenosine. ABX12681-ABX12698 represent antisense oligonucleotides  
CC for treating or preventing atopic diseases and neoplastic cell  
CC proliferation.  
XX  
SQ Sequence 19 BP; 2 A; 6 C; 6 G; 5 T; 0 other;  
Query Match 100.0%; Score 19; DB 25; Length 19;  
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Gaps 0;  
Matches 19; Conservative 0; Indels 0; Gaps 0;  
QY 1 CTGGGCCATCAGTCTCTG 19  
1 CTGGGCCATCAGTCTCTG 19  
DB 1 CTGGGCCATCAGTCTCTG 19  
RESULT 3  
AAF21266/c  
ID AAF21266 standard; DNA; 1068 BP.  
XX  
AC AAF21266;  
XX  
DT 14-MAR-2001 (first entry)  
XX  
DE Human low adenosine antisense oligonucleotide related sequence #2833.  
XX  
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
KW human; airway disorder; bronchoconstriction; lung inflammation;  
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
KW immunosuppressive; antiaesthetic; analgesic; hypotensive; cytostatic;  
KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
KW cancer; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200062736-A2.  
XX  
PD 26-OCT-2000.  
XX  
PF 24-MAR-2000; 2000WO-US08020.  
XX  
PR 06-APR-1999; 99US-0127958.  
XX  
PA (UYEC-) UNIV EAST CAROLINA.  
XX  
PI (NYCE/) NYCE J W.  
XX  
XX Myce JW;  
XX

DR WPI; 2000-679539/66.  
 XX Low adenosine (A) content antisense oligonucleotides which do not  
 PT trigger adenosine receptors during metabolism, useful e.g. for treating  
 PT cancers and respiratory obstructions -  
 XX  
 PS Disclosure; Page 1182; 1592pp; English.  
 XX  
 CC The present invention describes low adenosine (A) content antisense  
 CC oligonucleotides and compositions (I) comprising them. In the antisense  
 CC oligonucleotides the A is replaced by a 'universal' or alternative base.  
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 CC immunosuppressive, antiaesthetic, hypotensive and cytoskeletal activities.  
 CC The antisense oligonucleotides and (I) can be used to down-regulate the  
 CC expression and/or activity of target polypeptides associated with  
 CC lung/respiratory disorders and malignancies, such as stimulating and  
 CC activating peptide factors and transmitters, transcription factors,  
 CC immunoglobulin and antibodies, antibody receptors, cytokines and  
 CC chemokines, endogenously produced specific and non-specific enzymes,  
 CC binding proteins, adhesion molecules and their receptors, cytokine and  
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
 CC nervous system (CNS) and peripheral nervous and non-nervous system  
 CC receptors, CNS and peripheral nervous and non-nervous system peptide  
 CC transmitters, defensive, growth factors, vasoactive peptides and  
 CC receptors, binding proteins and malignancy associated proteins. The  
 CC antisense oligonucleotides may be used in this way to treat disorders  
 CC including respiratory obstruction (especially pulmonary obstruction  
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)  
 CC and/or surfactant hypoproduction which are associated with a disease or  
 CC condition selected from pulmonary vasoconstriction, inflammation,  
 CC allergies, asthma, impaired respiration, respiratory distress syndrome  
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
 CC fragments and antisense oligonucleotides used in the exemplification of  
 CC the present invention.  
 CC  
 XX Sequence 1068 BP; 231 A; 289 C; 243 G; 305 T; 0 other;  
 SO  
 Query Match 100.0%; Score 19; DB 21; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCCATCAGTCCTG 19  
 DB 105 CTGGGCCATCAGTCCTG 87  
 AAA35144/C  
 ID AAA35144 standard; DNA; 1068 BP.  
 AC AAA35144;  
 XX  
 DT 28-JUL-2000 (first entry)  
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:18.  
 XX  
 XX Human: adenosine receptor; low adenosine antisense oligonucleotide;  
 KM phosphothioate; impaired respiration; inflammation; allergy;  
 KM allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KM anti-allergic; antiaesthetic; cytoskeletal; analgesic; implanted airway;  
 KM lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KM respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KM pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KM cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200009525-A2.  
 XX  
 PD 24-FEB-2000.

XX  
 PF 03-AUG-1999; 99WO-US17712.  
 XX  
 PR 03-AUG-1998; 98US-0095212.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI NYCE JW;  
 XX  
 DR WPI; 2000-205971/18.  
 XX  
 PT New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
 PT cancers -  
 XX  
 PS Disclosure; Page 1102; 1343pp; English.  
 XX  
 CC The present invention describes a new composition comprising an  
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which  
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or  
 CC inflammation. The ON can have antiinflammatory, antiallergic,  
 CC antiaesthetic, cytoskeletal and analgesic activities. The compositions are  
 CC useful for the treatment of diseases associated with inflammation,  
 CC impaired airways, including lung disease and diseases whose secondary  
 CC effects afflict the lungs of a subject. They can be used for treating  
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,  
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic  
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,  
 CC carcinomas, and cancers which may metastasise to the lungs, including  
 CC breast and prostate cancer. The reduction of the adenosine content of  
 CC the ONs reduces side effects. The A-containing ONs break down with the  
 CC release of deoxyadenosine which activates adenosine receptors causing  
 CC bronchoconstriction and inflammation. AAA2313 to AAA3512 represent the  
 CC nucleotide sequences given in the sequence listing from the present  
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last  
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences  
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680  
 CC (AAA2323 to AAA3392) are specifically claimed ONs from the present  
 CC invention. N.B. Sequences given in the disclosure of the present  
 CC invention do not match up with their corresponding SEQ ID NO: sequences  
 CC given in the sequence listing.  
 CC  
 XX Sequence 1068 BP; 231 A; 289 C; 243 G; 305 T; 0 other;  
 SO  
 Query Match 100.0%; Score 19; DB 21; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCCATCAGTCCTG 19  
 DB 105 CTGGGCCATCAGTCCTG 87  
 AB197977/C  
 ID AB197977 standard; cDNA; 1068 BP.  
 AC AB197977;  
 XX  
 DT 18-FEB-2002 (first entry)  
 DE Non-endogenous human GPCR cDNA, SEQ ID NO: 474.  
 XX  
 XX Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;  
 KM constitutively activated GPCR; agonist; disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200177172-A2.  
 XX

```
PD 18-OCT-2001.
XX
PF 05-APR-2001; 2001WO-US11098.
XX
PR 07-APR-2000; 2000US-195747P.
XX
XX (AREN-) ARENA PHARM INC.
XX
PI Lehmann-Brulsma K, IJaw CW, Lin I;
XX
DR WPI; 2001-648759/74.
XX
DR P-PSDB; ABB56341.
XX
PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
PT disease treatment, comprises contacting candidate compounds with
PT versions of GPCRs -
XX
XX Example 2; Page 275; 394pp; English.
XX
CC The invention relates to G protein-coupled receptors (GPCRs) for which
CC the endogenous ligand has been identified. Non-endogenous
CC constitutively activated versions of known GPCRs are used in the
CC invention for the direct identification of candidate compounds as
CC receptor agonists, inverse agonists or partial agonists. Such
CC agonists are useful as therapeutic agents for diseases or disorders
CC associated with GPCRs. The present sequence encodes a non-endogenous
CC version of a known human GPCR.
XX
SQ Sequence 1068 BP; 232 A; 288 C; 244 G; 304 T; 0 other;

Query Match          100.0%; Score 19; DB 23; Length 1068;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTGTG 19
DB 105 CTGGGCCATCAGTGTCTGTG 87

RESULT 6
AB94340/C
ID AB94340 standard; DNA; 1068 BP.
XX
AC AB94340;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human CC chemokine receptor 3 (CCR3) encoding DNA.
XX
KW CC chemokine receptor-3; CCR3; antiallergic; antiinflammatory; human;
KW antiaesthetic; ophthalmological; dermatological; immunosuppressive;
KW antipruritic; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1068
FT FT /*tag= a
FT FT /product= "CCR3"

WO200192520-A1.
XX
PN 06-DEC-2001.
XX
PD 31-MAY-2001; 2001WO-EP06195.
XX
PR 01-JUN-2000; 2000GB-0013345.
XX
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Barnes AA, Frazer NJ, O'Shaughnessy CT, Wise A;
XX
DR WPI; 2002-114347/15.
```

```
DR P-PSDB; ABB07240.
XX
PT Modified CC chemokine receptor-3 useful for identifying modulators of
PT eotaxin-mediated CCR3 receptor for treating allergic and inflammatory
PT disorders, comprises modifications to stabilize or enhance surface
PT expression -
XX
XX Disclosure; Page 25-26; 29pp; English.
XX
CC The invention relates to a CC chemokine receptor-3 (CCR3) modified to
CC stabilize or enhance expression of the receptor in a cell membrane.
CC Assays for investigating properties of the CCR3 receptor are useful for
CC the identification of modulators of eotaxin-mediated CCR3 receptor
CC activity. The identified modulators are useful in the treatment of
CC prophylaxis of allergic or inflammatory disorders which are responsive to
CC regulation of CCR3 receptor activity. The agents are also useful in the
CC treatment of allergy or asthma as well as ophthalmological, inflammatory,
CC gastrointestinal, dermatological, respiratory or pruritic disorders. The
CC agents are useful for treating conjunctivitis, inflammatory bowel disease,
CC eczema, allergic rhinitis, nasal polyposis, atopic dermatitis and
CC pruritis, chronic obstructive pulmonary disease (COPD) and other lung
CC disorders and immune disease. The present sequence represents the human
CC CCR3 receptor encoding DNA.
XX
SQ Sequence 1068 BP; 231 A; 289 C; 243 G; 305 T; 0 other;

Query Match          100.0%; Score 19; DB 24; Length 1068;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTGTG 19
DB 105 CTGGGCCATCAGTGTCTGTG 87

RESULT 7
AAD25222/C
ID AAD25222 standard; cDNA; 1068 BP.
XX
AC AAD25222;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human chemokine (C-C motif) receptor 3 (CCR3) cDNA.
XX
KW Human; chemokine (C-C motif) receptor 3; CCR3 gene; haplotyping;
KW genotyping; type IV hypersensitivity reaction; HIV-1; gene therapy;
KW human immunodeficiency virus 1; single nucleotide polymorphism; SNP;
KW chromosome 3p21.3; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace (51, C)
FT FT /*tag= a
FT FT /standard name= "Single nucleotide polymorphism (SNP)"
FT FT /note= "Polymorphic site (PS) 3"
FT FT 1..1068
FT FT /*tag= b
FT FT /product= "Human CCR3 protein"
FT FT /replace (1052, C)
FT FT /*tag= c
FT FT /standard name= "Single nucleotide polymorphism (SNP)"
FT FT /note= "Polymorphic site (PS) 4"

WO200187908-A2.
XX
PN 22-NOV-2001.
XX
PD 18-MAY-2001; 2001WO-US16278.
XX
PR 18-MAY-2000; 2000US-205191P.
XX
```

PA (GENA-) GENAISSANCE PHARM INC.  
 XX  
 PI Choi JY, Kazemi A, Koshy B;  
 XX  
 DR WPI; 2002-055681/07.  
 DR P-PSDB; AAE15320.  
 XX  
 PT Isolated polymorphic variants of chemokine (C-C motif) receptor 3  
 PT (CCR3) gene useful for studying function of CCR3, expressing the CCR3  
 PT protein and to screen drugs to treat CCR3 activity-related diseases -  
 XX  
 PS Claim 25a; Fig 2; 53pp; English.  
 XX  
 CC The invention relates to genetic variants of human chemokine (C-C motif)  
 CC receptor 3 (CCR3) gene. The invention also relates to compositions and  
 CC methods for haplotyping and/or genotyping the CCR3 gene in an individual.  
 CC Polynucleotides of the invention are useful for studying the expression  
 CC and function of CCR3 and in expressing CCR3 proteins for use in screening  
 CC candidate drugs to treat diseases related to CCR3 activity. They are also  
 CC used in gene therapy. The polymorphism and haplotype data is useful for  
 CC validating whether CCR3 is a suitable target for drugs to treat type IV  
 CC hypersensitivity reactions and human immunodeficiency virus (HIV)-1,  
 CC screening for such drugs and reducing bias cells in clinical trials of  
 CC such drugs. The genotyping method is useful for determining whether an  
 CC individual has one haplotype or haplotype pairs. The haplotyping method  
 CC is useful for improving the efficiency and outcome of several steps in  
 CC the discovery and development of drugs for treating diseases associated  
 CC with CCR3 activity such as type IV hypersensitivity reactions and HIV-1.  
 CC The present sequence is human CCR3 cDNA. The CCR3 gene is located on  
 CC chromosome 3p21.3.  
 XX  
 SQ Sequence 1068 BP; 231 A; 289 C; 243 G; 305 T; 0 other;  
 XX  
 Query Match 100.0%; Score 19; DB 24; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 CTGGGCCATCAGTGTCTCTG 19  
 |||||  
 DB 105 CTGGGCCATCAGTGTCTCTG 87  
 XX  
 RESULT 8  
 AAT79096/c  
 ID AAT79096 standard; cDNA; 1071 BP.  
 XX  
 AC AAT79096;  
 XX  
 DT 13-MAR-1998 (first entry)  
 XX  
 DE Human CCR3 chemokine receptor coding sequence.  
 XX  
 KW CCR3 chemokine; mouse; primer; PCR; amplification; antagonist; human;  
 KW abnormal physiology; development; anti-viral; probe; hybridisation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT 1..1071  
 FT /\*tag= a  
 FT /product= CCR3 chemokine receptor  
 XX  
 MO9721812-A2.  
 XX  
 PD 19-JUN-1997.  
 XX  
 PF 05-DEC-1996; 96MO-US19139.  
 XX  
 PR 08-DEC-1995; 95US-0567882.  
 XX  
 PA (SCHE) SCHERING CORP.  
 XX  
 PI Daireghi DJ, Hara T, Miyajima A, Schall TJ, Wang W;

PI Yoshimura A;  
 XX  
 DR WPI; 1997-332784/30.  
 DR P-PSDB; AAW25943.  
 XX  
 PT New isolated chemokine CCR3 and chemokine receptor CCR3 - used to  
 PT develop products useful for the diagnosis and treatment of  
 PT conditions associated with abnormal physiology or development  
 XX  
 PS Claim 15; Page 60-62; 73pp; English.  
 XX  
 CC This is the nucleotide sequence encoding a novel CCR3 chemokine  
 CC receptor isolated from a Th0-activated human T-cell cDNA library  
 CC using the sequence amplified by primers AAT79097 and AAT79098 as a  
 CC probe. The encoded protein can be used to screen for (ant)agonists that  
 CC bind to the novel CCR3 chemokines (AAW25941 and AAW25942). These  
 CC (ant)agonists are useful in the treatment of conditions associated with  
 CC abnormal physiology or development.  
 XX  
 SQ Sequence 1071 BP; 231 A; 292 C; 242 G; 306 T; 0 other;  
 XX  
 Query Match 100.0%; Score 19; DB 18; Length 1071;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 CTGGGCCATCAGTGTCTCTG 19  
 |||||  
 DB 105 CTGGGCCATCAGTGTCTCTG 87  
 XX  
 RESULT 9  
 AAT3336/c  
 ID AAT3336 standard; DNA; 1116 BP.  
 XX  
 AC AAT3336;  
 XX  
 DT 15-NOV-1996 (first entry)  
 XX  
 DE CC-chemokine receptor 3 consensus DNA sequence.  
 XX  
 KW CC-chemokine receptor 3; CCR3; Eos-12; inhibitor; antisense;  
 KW antiinflammatory; eosinophil; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO9622371-A2.  
 XX  
 PD 25-JUL-1996.  
 XX  
 PF 19-JAN-1996; 96MO-US00608.  
 XX  
 PR 19-JAN-1995; 95US-0375199.  
 XX  
 PA (BGM) BRIGHAM & WOMENS HOSPITAL.  
 PA (CHIL-) CHILDRENS MEDICAL CENT.  
 PA (LEUK-) LEUKOSITE INC.  
 XX  
 PI Gerard CJ, Gerard NP, Mackay CR, Ponath PD, Post TW;  
 PI Qin S;  
 XX  
 DR WPI; 1996-354528/35.  
 DR P-PSDB; AAW03378.  
 XX  
 PT Mammalian chemokine receptor-3 and related nucleic acids - useful to  
 PT identify receptor inhibitors to treat inflammatory disease, e.g.  
 PT autoimmune disorders, certain cancers, etc.  
 XX  
 PS Claim 1; Page 114-115; 153pp; English.  
 XX  
 CC A consensus DNA sequence (AAT3336) codes for a novel human receptor  
 CC (AAW03378), designated Eos 12 or C-C chemokine receptor 2 (CCR-3).  
 CC It was deduced by comparing a genomic clone (AAT3334) and a cDNA  
 CC clone (AAT3335) coding for CCR-3 protein (AAW03376 and AAW03377)

CC having slightly different sequences. Initial sequence information  
CC revealed 2 regions in which the cDNA sequence appeared to be shifted in  
CC frame, resulting in 2 sets of 4 contiguous amino acid differences  
CC in the predicted proteins. Further sequence analysis revealed only  
CC a single difference between the 2 open reading frames, the genomic  
CC clone coding for threonine at position 276 and the cDNA clone for  
CC serine.

XX  
SQ Sequence 1116 BP; 246 A; 286 C; 257 G; 306 T; 21 other;

Query Match 100.0%; Score 19; DB 17; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTCTG 19  
|||  
Db 119 CTGGGCCATCAGTGTCTCTG 101

RESULT 10  
AAV07404/C  
ID AAV07404 standard; DNA; 1116 BP.  
XX  
AC AAV07404;  
XX  
DT 28-SEP-1998 (first entry)  
XX  
DE Human C-C chemokine receptor 3 consensus nucleic acid.  
XX  
KW C-C chemokine receptor 3; CCR3; Bos L2; human;  
KW G protein-coupled receptor; leukocyte; antibody; antagonist;  
KW inflammation; allergy; asthma; graft rejection; infection;  
KW autoimmune disease; drug screening; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 15..1082  
FT /\*tag= a  
FT  
FN WO9814480-A1.  
XX  
PD 09-APR-1998.  
XX  
PF 24-SEP-1997; 97WO-US17103.  
XX  
PR 30-SEP-1996; 96US-0720565.  
XX  
PA (LEUK-) LEUKOSITE INC.  
XX  
PI Mackay CR, Ponath PD;  
XX  
DR WPI; 1998-286418/25.  
DR P-PSDB; AAW51746.  
XX  
PT Antibodies to chemokine receptor-3 protein - useful for diagnosis  
PT and treatment of inflammatory conditions, e.g. allergy, asthma,  
PT autoimmune disease, graft rejection or cancer  
XX  
PS Disclosure; Page 137-138; 185pp; English.  
XX  
CC This is a consensus sequence constructed by alignment of a genomic  
CC DNA sequence (see AAV07402) and a cDNA clone (see AAV07403) coding for  
CC novel human C-C chemokine receptor 3, also designated CCR-3, CCR3  
CC or Bos L2, that binds and mediates chemotaxis in response to  
CC chemokines such as eotaxin, RANTES and MCP-3. Sequence comparison  
CC revealed 2 regions in the cDNA sequence that appeared to be shifted  
CC in frame, resulting from an insertion of a base followed by the  
CC deletion of a base, or the deletion of a base followed by the  
CC insertion of a base. These alterations resulted in 4 contiguous  
CC amino acid differences in the predicted proteins (see AAW51744 and  
CC AAW51745) at positions 263-266 and 276-279, respectively. In  
CC addition, the genomic clone codes for threonine (ACG) at

CC position 276 and the cDNA clone for serine (AGC). The predicted  
CC amino acid sequence from the consensus is given in AAW51747.  
CC CCR-3 nucleic acids, polypeptides, antibodies, agonists and  
CC antagonists are useful for diagnosis and treatment of  
CC inflammatory conditions, autoimmune diseases and infections.  
XX  
SQ Sequence 1116 BP; 246 A; 287 C; 257 G; 306 T; 20 other;

Query Match 100.0%; Score 19; DB 19; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTGTCTCTG 19  
|||  
Db 119 CTGGGCCATCAGTGTCTCTG 101

RESULT 11  
AAT31335/C  
ID AAT31335 standard; cDNA; 1193 BP.  
XX  
AC AAT31335;  
XX  
DT 15-NOV-1996 (first entry)  
XX  
DE CC-chemokine receptor 3 cDNA clone.  
XX  
KW CC-chemokine receptor 3; CCR-3; Bos-L2; inhibitor; antisense;  
KW antiinflammatory; eosinophil; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 92..1159  
FT /\*tag= a  
FT variation 918..919  
FT /\*tag= b  
FT /note= "CCR-3 cDNA clone has GC at positions  
FT 918-919, coding for serine (AGC) at  
FT position 276; a genomic clone has CG at  
FT these positions, coding for threonine  
FT (AGC)"  
FT  
FN WO9622371-A2.  
XX  
PD 25-JUL-1996.  
XX  
PF 19-JAN-1996; 96WO-US00608.  
XX  
PR 19-JAN-1995; 95US-0375199.  
XX  
PA (BGMH) BRIGHAM & WOMEN'S HOSPITAL.  
PA (CHIL-) CHILDRENS MEDICAL CENT.  
PA (LEUK-) LEUKOSITE INC.  
XX  
PI Gerard CJ, Gerard NP, Mackay CR, Ponath PD, Post TW;  
PI Qin S;  
XX  
DR WPI; 1996-354528/35.  
DR P-PSDB; AAW03377.  
XX  
PT Mammalian chemokine receptor-3 and related nucleic acids - useful to  
PT identify receptor inhibitors to treat inflammatory disease, e.g.  
PT autoimmune disorders, certain cancers, etc.  
XX  
PS Claim 1; Page 111-113; 153pp; English.  
XX  
CC A genomic DNA clone (AAT31335) codes for a novel receptor (AAW03377),  
CC designated Bos L2 or C-C chemokine receptor 3 (CCR-3), involved  
CC in leukocyte migration associated with inflammation. It was  
CC isolated from a human library constructed from eosinophils obtd.  
CC from a patient with hyper-eosinophilic syndrome using a probe  
CC (p4 cDNA) encoding the MIP-1alpha/RANTES receptor. A CCR-3

CC genomic clone (AA11334) was also isolated, and a consensus sequence  
 CC is given in AA11336. The cDNA and genomic clones can be used for  
 CC the prodn. of recombinant CR-3 in host cells, or to design  
 CC antisense sequences useful for treating inflammatory disease.  
 XX  
 SQ Sequence 1193 BP; 274 A; 310 C; 275 G; 334 T; 0 other;  
 Query Match 100.0%; Score 19; DB 17; Length 1193;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCCATCAGTGTCTG 19  
 |||||  
 Db 196 CTGGGCCATCAGTGTCTG 178  
 RESULT 12  
 AA07403/c  
 ID AA07403 standard; cDNA; 1193 BP.  
 XX  
 AC AA07403;  
 XX  
 DT 28-SEP-1998 (first entry)  
 XX  
 DE Human C-C chemokine receptor 3 cDNA.  
 XX  
 KM C-C chemokine receptor 3; CR-3; CCR3; Eos I2; human;  
 KM G protein-coupled receptor; leukocyte; antibody; antagonist;  
 KM inflammation; allergy; asthma; graft rejection; infection;  
 KM autoimmune disease; drug screening; therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 92..1159  
 FT /\*tag= a  
 FT  
 PN MO9814480-A1.  
 XX  
 PD 09-APR-1998.  
 XX  
 PF 24-SEP-1997; 97MO-US17103.  
 XX  
 PR 30-SEP-1996; 96US-0720565.  
 XX  
 PA (LEUK-) LEUKOSITTE INC.  
 XX  
 PI Mackay CR, Ponath PD;  
 XX  
 DR WPI: 1998-286418/25.  
 DR P-PSDB; AAW51745.  
 XX  
 PT Antibodies to chemokine receptor-3 protein - useful for diagnosis  
 PT and treatment of inflammatory conditions, e.g. allergy, asthma,  
 PT autoimmune disease, graft rejection or cancer  
 XX  
 PS Example 8; Page 134-136; 185pp; English.  
 XX  
 CC This cDNA codes for novel human C-C chemokine receptor 3 (see  
 CC AAW51745), also designated CR-3, CCR3 or Eos I2, that binds and  
 CC mediates chemotaxis in response to chemokines such as eotaxin,  
 CC RANTES and MCP-3. The cDNA was isolated from a human eosinophil  
 CC cDNA library constructed from eosinophils obtained from a patient  
 CC with hyperesinophilic syndrome, and using CR-1 cDNA as probe. A  
 CC genomic DNA sequence (see AA07402) is also provided as well as a  
 CC consensus sequence (see AA07404) for CR-3. The invention relates  
 CC to isolated and/or recombinant nucleic acids encoding CR-3,  
 CC isolated or recombinant CR-3 polypeptides, recombinant nucleic  
 CC acid constructs, host cells useful for production of recombinant  
 CC CR-3 proteins, to antibodies reactive with the receptors, and to  
 CC methods of using these products to identify ligands, antagonists  
 CC and agonists of receptor function. Inhibitors of CR-3 can be used  
 CC to treat: inflammatory or allergic diseases and conditions,

CC including respiratory allergic diseases such as asthma, allergic  
 CC rhinitis, hypersensitivity lung disease, hypersensitivity  
 CC pneumonia, eosinophilic pneumonia (e.g. Loeffler's syndrome,  
 CC chronic eosinophilic pneumonia, interstitial lung disease (ILD)  
 CC e.g. idiopathic pulmonary fibrosis or ILD associated with  
 CC rheumatoid arthritis, systemic lupus erythematosus, ankylosing  
 CC spondylitis, systemic sclerosis, Sjogren's syndrome, polymyositis  
 CC or dermatomyositis), systemic anaphylaxis or hypersensitivity  
 CC responses, drug allergy, insect sting allergy, inflammatory bowel  
 CC disease, such as Crohn's disease and ulcerative colitis,  
 CC spondyloarthritis, scleroderma, psoriasis, inflammatory  
 CC dermatosis such as dermatitis, eczema, atopic dermatitis,  
 CC allergic contact dermatitis, urticaria, vasculitis (e.g. necrotizing,  
 CC cutaneous and hypersensitivity vasculitis); eosinophilic myositis  
 CC and eosinophilic fasciitis; autoimmune diseases such as rheumatoid  
 CC arthritis, psoriatic arthritis, multiple sclerosis, systemic lupus  
 CC erythematosus, myasthenia gravis, juvenile onset diabetes;  
 CC glomerulonephritis, autoimmune thyroiditis and Behcet's disease;  
 CC graft rejection, including allograft rejection or graft-versus-host  
 CC disease; cancers with leukocyte infiltration of the skin or organs;  
 CC and also reperfusion injury, atherosclerosis, certain hematologic  
 CC malignancies, septic shock and endotoxic shock. Promoters of CR-3  
 CC function can be used for treating: immunosuppression e.g. in AIDS  
 CC patients or individuals undergoing radiation therapy, chemotherapy,  
 CC therapy for autoimmune disease or other drug therapy, and  
 CC immunosuppression due congenital deficiency in receptor function or  
 CC other causes; and infectious diseases such as parasitic diseases,  
 CC including helminth infections, such as nematodes (round worms).  
 CC The agents can also be used for detection and diagnosis.  
 XX  
 SQ Sequence 1193 BP; 274 A; 310 C; 275 G; 334 T; 0 other;  
 Query Match 100.0%; Score 19; DB 19; Length 1193;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCCATCAGTGTCTG 19  
 |||||  
 Db 196 CTGGGCCATCAGTGTCTG 178  
 RESULT 13  
 AAF21267/c  
 ID AAF21267 standard; DNA; 1201 BP.  
 XX  
 AC AAF21267;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Human low adenovine antisense oligonucleotide related sequence #2834.  
 XX  
 KM Low adenovine antisense oligonucleotide; phosphorothioate; allergy;  
 KM human; airway disorder; bronchoconstriction; lung inflammation;  
 KM surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 KM immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;  
 KM respiratory obstruction; pulmonary; impeded respiration;  
 KM surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 KM respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 KM pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 KM chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 KM cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200062736-A2.  
 XX  
 PD 26-OCT-2000.  
 XX  
 PF 24-MAR-2000; 2000MO-US08020.  
 XX  
 PR 06-APR-1999; 99US-0127958.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.

PA (NYCE/) NYCE J W.  
XX  
PI Nyce JW;  
XX  
DR WP1: 2000-679539/66.  
XX  
PT Low adenosine (A) content antisense oligonucleotides which do not  
PT trigger adenosine receptors during metabolism, useful e.g. for treating  
PT cancers and respiratory obstructions -  
XX  
XX Disclosure; Page 1182-1183; 1592pp; English.  
PS  
XX  
XX The present invention describes low adenosine (A) content antisense  
CC oligonucleotides and compositions (I) comprising them. In the antisense  
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.  
CC The antisense oligonucleotides and (I) can be used to down-regulate the  
CC expression and/or activity of target polypeptides associated with  
CC lung/respiratory disorders and malignancies, such as stimulating and  
CC activating peptide factors and neurotransmitters, transcription factors,  
CC immunoglobulins and antibodies, antibody receptors, cytokines and  
CC chemokines, endogenously produced specific and non-specific enzymes,  
CC binding proteins, adhesion molecules and their receptors, cytokine and  
CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
CC nervous system (CNS) and peripheral nervous and non-nervous system  
CC receptors, CNS and peripheral nervous and non-nervous system peptide  
CC transmitters, defensins, growth factors, vasoactive peptides and  
CC receptors, binding proteins and malignancy associated proteins. The  
CC antisense oligonucleotides may be used in this way to treat disorders  
CC including respiratory obstruction (especially pulmonary obstruction  
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)  
CC and/or surfactant hypoproduction which are associated with a disease or  
CC condition selected from pulmonary vasoconstriction, inflammation,  
CC allergies, asthma, impaired respiration, respiratory distress syndrome  
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
CC and/or cancer. AAPl8434 to AAPl21543 represent human polynucleotide  
CC fragments and antisense oligonucleotides used in the exemplification of  
CC the present invention.  
CC  
XX  
SQ Sequence 1201 BP: 278 A; 320 C; 267 G; 336 T; 0 other;

Query Match 100.0%; Score 19; DB 21; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCATCATGCTCTCG 19  
||| ||||||| ||||||| |||

DB 136 CTGGGCATCATGCTCTCG 118

RESULT 14  
ID AAA35145/c  
XX AAA35145 standard; DNA; 1201 BP.  
XX  
XX AAA35145;  
XX  
DT 28-JUL-2000 (first entry)

DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:19.

XX  
XX Human, adenosine receptor; low adenosine antisense oligonucleotide;  
KM phosphorothioate; impaired respiration; inflammation; allergy;  
KM allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
KM antiallergic; antiasthmatic; cytosstatic; analgesic; impaired airway;  
KM lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
KM respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
KM pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
KM cancer; leukaemia; lymphoma; carcinoma; metastasis; BG.

XX Homo sapiens  
XS

XX	PN	WO200009525-A2.
XX	PD	
XX	PD	24-FEB-2000.
XX	PF	03-AUG-1999; 99WO-US17712.
XX	PR	03-AUG-1998; 98US-0095212.
XX	PR	(UYEC-) UNIV EAST CAROLINA.
XX	PI	Myce JW;
XX	PI	WPI; 2000-205971/18.
XX	DR	
XX	PT	New antinease oligonucleotides useful for treating e.g. pulmonary
XX	PT	vasoconstriction, inflammation, allergies, asthma, hypertension,
XX	PT	bronchitis, emphysema, respiratory distress syndrome, ischemia or
XX	PT	cancers -
XX	XX	
XX	XX	Disclosure; Page 1102, 1343pp; English.
XX	XX	
XX	XX	The present invention describes a new composition comprising an
XX	XX	antinease oligonucleotide (ON) with low adenosine (up to 15%), which
XX	XX	targets nucleic acids involved in bronchoconstriction, allergies, and/or
XX	XX	inflammation. The ON can have antiinflammatory, antiallergic,
XX	XX	antiasthmatic, cytosstatic and analgesic activities. The compositions are
XX	XX	useful for the treatment of diseases associated with inflammation,
XX	XX	impaired airways, including lung disease and diseases whose secondary
XX	XX	effects afflict the lungs of a subject. They can be used for treating
XX	XX	e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
XX	XX	asthma, impeded respiration, respiratory distress syndrome, pain, cystic
XX	XX	fibrosis, pulmonary hypertension, emphysema, chronic obstructive
XX	XX	pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
XX	XX	carcinomas, and cancers which may metastasise to the lungs, including
XX	XX	breast and prostate cancer. The reduction of the adenosine content of
XX	XX	the ONs reduces side effects. The A-containing ONs break down with the
XX	XX	release of deoxyadenosine which activates adenosine receptors causing
XX	XX	bronchoconstriction and inflammation. AAA33313 to AAA35312 represent the
XX	XX	nucleotide sequences given in the sequence listing from the present
XX	XX	invention, which correspond to SEQ ID NO:1 to 2815, and then the last
XX	XX	185 sequences are also called SEQ ID NO:1 to 185, and then the sequences
XX	XX	CC differ from the previously named sequences. SEQ ID NO:11 to 1680
XX	XX	(AAA33323 to AAA33992) are specifically claimed ONs from the present
XX	XX	invention. N.B. Sequences given in the disclosure of the present
XX	XX	invention do not match up with their corresponding SEQ ID NO: sequences
XX	XX	given in the sequence listing.
XX	XX	
XX	XX	Sequence 1201 BP; 278 A; 320 C; 267 G; 336 T; 0 other;
XX	XX	
XX	XX	Query Match 100.0%; Score 19; DB 21; Length 1201;
XX	XX	Best Local Similarity 100.0%; Pred. No. 5.9;
XX	XX	Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	XX	
XX	XX	1 CTGGGCATCATGTCCTCTG 19
XX	XX	
XX	XX	136 CTGGGCATCATGTCCTCTG 118
XX	XX	
XX	XX	RESULT 15
XX	XX	ABK84282/C
XX	XX	ID ABK84282 standard; cDNA; 1201 BP.
XX	XX	ABK84282;
XX	XX	
XX	XX	14-AUG-2002 (first entry)
XX	XX	
XX	XX	Human cDNA differentially expressed in granulocytic cells #853.
XX	XX	
XX	XX	Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX	XX	viral infection; parasitic infection; protozoal infection;
XX	XX	fungal infection; sterile inflammatory disease; psoriasis;
XX	XX	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;



KM cardiac reperfusion injury; renal reperfusion injury; ARDS;  
 KM adult respiratory distress syndrome; inflammatory bowel disease;  
 KM Crohn's disease; ulcerative colitis; peridontal disease;  
 KM granulocyte activation; chronic inflammation; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO20028999-A2.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 03-OCT-2001; 2001MO-US30821.  
 XX  
 PR 03-OCT-2000; 2000US-237189P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
 PI WPI; 2002-435328/46.  
 DR  
 XX  
 XX  
 PT Detecting granulocyte activation by detecting differential expression  
 PT of genes associated with granulocyte activation, which serves as  
 PT diagnostic markers that is useful for monitoring disease states and  
 PT drug toxicity -  
 XX  
 PS Claim 1; SEQ ID NO 853; 114pp; English.  
 CC The invention relates to detecting (M1) granulocyte (GC) activation  
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
 CC DNA chip analysis as given in the specification, and comparing  
 CC the expression level to an expression level in an unactivated  
 CC GC, where differential expression of Gs is indicative of GCA.  
 CC Also included are modulating (M2) Gs by contacting GC with an agent  
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)  
 CC for an agent capable of modulating GCA or an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease using the  
 CC gene expression profile; (3) detecting (M4) an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease, by detecting the  
 CC level of expression in a sample of the tissue of gene(s) from Gs, where  
 CC the level of expression of the gene is indicative of inflammation;  
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,  
 CC an allergic response in a subject, exposure of a subject to a pathogen  
 CC or sterile inflammatory disease, by contacting a tissue having  
 CC inflammation with an agent that modulates the expression of gene(s)  
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for  
 CC modulating GCA; M3 is useful for screening an agent capable of modulating  
 CC GCA preferably in an inflammation in a tissue; M4 is useful for  
 CC detecting an inflammation (especially chronic) in a tissue, an allergic  
 CC response in a subject, exposure of a subject to a pathogen or sterile  
 CC inflammatory disease (e.g., psoriasis, rheumatoid arthritis,  
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal  
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,  
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,  
 CC periodontal disease; also bacterial infection, viral infection,  
 CC parasitic infection, protozoal infection, fungal infection and MS is  
 CC useful for treating one of the above conditions. The present  
 CC sequence represents a gene differentially expressed in granulocytes.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 1201 BP; 278 A; 320 C; 267 G; 336 T; 0 other;  
 XX  
 XX  
 Query Match 100.0%; Score 19; DB 24; Length 1201;  
 Best Local Similarity 100.0%; Pred. No. 5.9; Indels 0; Gaps 0;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGGCATCAGTGTCTG 19  
 ||||||||||||||||

DB 136 CTGGGGCATCAGTGTCTG 118  
 RESULT 16  
 ACAS6487/C  
 ID ACAS6487 standard; cDNA; 1201 BP.  
 XX  
 AC ACAS6487;  
 XX  
 DT 06-JUN-2003 (first entry)  
 XX  
 DE Human signalling pathway polynucleotide probe SEQ ID NO 1085.  
 XX  
 KM Human; probe; ss; array element; Parkinson's disease;  
 KM signalling pathway population; cancer; adenocarcinoma; leukemia;  
 KM immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6500938-B1.  
 XX  
 PD 31-DEC-2002.  
 XX  
 PF 30-JAN-1998; 98US-0016434.  
 XX  
 PR 30-JAN-1998; 98US-0016434.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 PI Au-Young J, Sellhammer JU;  
 PI WPI; 2003-352189/33.  
 DR  
 XX  
 XX  
 PT Combination of polynucleotide probes, useful as array elements in a  
 PT microarray for monitoring the expression of a number of target  
 PT polynucleotides -  
 XX  
 PS Claim 1; SEQ ID NO 1085; 65pp; English.  
 CC The invention relates to a combination which comprises a number of  
 CC polynucleotide probes comprising a sequence selected from one of the 1490  
 CC sequences mentioned in the specification. The combination is useful as an  
 CC array element in a microarray for monitoring the expression of a number  
 CC of target polynucleotides. The microarray is particularly useful in the  
 CC diagnosis and treatment of cancer and immunopathology and neuropathology.  
 CC The microarray is useful in diagnostics and treatment regimens, drug  
 CC discovery and development, toxicological and carcinogenicity studies,  
 CC forensics and pharmacogenomics. The microarray is also useful for  
 CC monitoring progression of diseases and for developing sophisticated  
 CC profiles for the effects of currently available therapeutic drugs. The  
 CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs  
 CC and genomic fragments and in research and diagnostic applications. The  
 CC array can detect changes in expression in a large number of genes coding  
 CC for different signalling pathway populations which can be used to diagnose  
 CC various diseases including cancer e.g. adenocarcinoma and leukemia,  
 CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease  
 CC and Parkinson's disease. The present sequence represents a polynucleotide  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from USPTO  
 CC at seqdata.uspto.gov/sequence.html?DocId=06500938B1.  
 CC  
 SQ Sequence 1201 BP; 278 A; 320 C; 267 G; 336 T; 0 other;  
 XX  
 XX  
 Query Match 100.0%; Score 19; DB 25; Length 1201;  
 Best Local Similarity 100.0%; Pred. No. 5.9; Indels 0; Gaps 0;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGGCATCAGTGTCTG 19  
 ||||||||||||||||  
 DB 136 CTGGGGCATCAGTGTCTG 118



XX Sequence 1689 BP; 431 A; 416 C; 344 G; 497 T; 1 other;  
 SQ Query Match 100.0%; Score 19; DB 17; Length 1689;  
 Best Local Similarity 100.0%; Pred. No. 6.2;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
 |||||  
 DB 285 CTGGGCCATCAGTCTCTG 267

RESULT 19  
 AAT58783/C  
 ID AAT58783 standard; DNA; 1689 BP.  
 XX  
 AC AAT58783;  
 DT 30-SEP-1997 (first entry)  
 XX  
 DE Human C-C chemokine receptor 3 DNA.

Human, ectaxin; eosinophil; chemottractant; stimulation;  
 accumulation; attraction; chemotaxis; diagnosis; prevention;  
 treatment; disease; inflammation; allergy; asthma; rhinitis;  
 hypersensitivity; lung; pneumonia; Loeffler's; syndrome;  
 interstitial; ILD; idiopathic pulmonary fibrosis;  
 rheumatoid arthritis; systemic; lupus erythematosus; SLE;  
 ankylosing spondylitis; sclerosis; Sjorgen's; polymyositis;  
 dermatomyositis; bowel; anaphylaxis; drug; penicillin;  
 cephalosporin; insect sting; Crohn's; ulcerative colitis;  
 spondyloarthritis; scleroderma; psoriasis; dermatosis;  
 dermatitis; eczema; atopic; urticaria; necrotic; cutaneous;  
 vasculitis; myositis; fascitis; multiple sclerosis;  
 myasthenia gravis; juvenile onset diabetes; glomerulonephritis;  
 autoimmunity; thyroiditis; Bechet's; graft; rejection;  
 transplantation; allograft; graft versus host; cancer;  
 leukocyte infiltration; reperfusion injury; atherosclerosis;  
 haematologic malignancy; septic; endotoxic; shock;  
 polymyositis; dermatomyositis; immunosuppression; immunodeficiency;  
 AIDS; radiation therapy; chemotherapy; autoimmune; corticosteroid;  
 C-C chemokine receptor 3; CRK3; ss.

XX Homo sapiens.  
 OS  
 XX  
 FT Key Location/Qualifiers  
 CDS 181..1248  
 FT /\*tag= a  
 FT /product= CRK3

MO9700960-A1.  
 XX  
 PD 09-JAN-1997.  
 XX  
 PF 21-JUN-1996; 96WO-US10723.  
 XX  
 PR 23-JUN-1995; 95US-0494093.  
 XX  
 XX (LEUK-) LEUKOSITE INC.  
 XX  
 PI Mackay C, Newman W, Ponath PD, Qin S, Ringle DJ;  
 DR WPI; 1997-087387/08.  
 XX P-PSDB; AAM10100.  
 XX  
 PS New isolated human ectaxin gene - used to develop prods. for the  
 PT diagnosis and treatment of e.g. inflammation, allergies, auto-immune  
 PT disease, infections and tumours  
 XX  
 PS Example 7; Page 97; 130pp; English.  
 XX  
 CC The present sequence encodes human C-C chemokine receptor 3 (CRK3),  
 CC to which human ectaxin (hE), an eosinophil specific chemottractant

CC capable of stimulating eosinophil accumulation and/or attracting  
 CC eosinophils (including chemotaxis), binds.  
 CC hE can be used to develop products for the diagnosis, prevention or  
 CC treatment of hE associated diseases or conditions. The products can  
 CC be used to treat inflammatory or allergic diseases and conditions,  
 CC including respiratory allergic diseases (e.g. asthma, allergic  
 CC rhinitis, hypersensitivity lung diseases or pneumonia),  
 CC eosinophilic pneumonias such as Loeffler's syndrome and chronic  
 CC eosinophilic pneumonia, interstitial lung diseases (ILD) such as  
 CC idiopathic pulmonary fibrosis or ILD associated with rheumatoid  
 CC arthritis, systemic lupus erythematosus (SLE), ankylosing  
 CC spondylitis, systemic sclerosis, Sjorgen's syndrome, polymyositis  
 CC or dermatomyositis), systemic anaphylaxis or hypersensitivity  
 CC responses, drug allergies (e.g. to penicillin and cephalosporins),  
 CC insect sting allergies, inflammatory bowel diseases (e.g. Crohn's  
 CC disease and ulcerative colitis), spondyloarthritis,  
 CC scleroderma, psoriasis and inflammatory dermatoses (e.g.  
 CC dermatitis, eczema, atopic dermatitis, allergic contact dermatitis,  
 CC urticaria and necrotizing, cutaneous and hypersensitivity  
 CC vasculitis), eosinophilic myositis and fascitis, multiple  
 CC sclerosis, SLE, myasthenia gravis, juvenile onset diabetes,  
 CC glomerulonephritis, autoimmune thyroiditis, Bechet's disease, graft  
 CC rejection (e.g. in transplantation) including allograft rejection or  
 CC graft versus host disease and cancers with leukocyte infiltration  
 CC of the skin or organs. The products can also be used to treat other  
 CC diseases or conditions regulating the inhibition of undesirable  
 CC inflammatory responses, including reperfusion injury,  
 CC atherosclerosis, certain haematologic malignancies, cytokine  
 CC induced toxicity (e.g. septic or endotoxic shock), polymyositis,  
 CC dermatomyositis, immunosuppression (e.g. in individuals with  
 CC immunodeficiency syndromes such as AIDS, undergoing radiation  
 CC therapy, chemotherapy, therapy for autoimmune disease or other drug  
 CC therapy, such as corticosteroid therapy, which causes  
 CC immunosuppression), immunosuppression due to (e.g. congenital)  
 CC deficiency (e.g. in ectaxin) or infectious diseases such as parasitic  
 CC diseases.  
 CC Degenerate primers based on the guinea pig ectaxin amino acid  
 CC sequence were used for the reverse transcriptase polymerase chain  
 CC reaction (RT-PCR) amplification of RNA isolated from inflamed,  
 CC eosinophilic lung tissue obtained from Balb/c mice sensitised to  
 CC ovalbumin. The amplification product was used as a probe to screen  
 CC a human genomic library in vector EMBL3 SP6/T7 to obtain the hE  
 CC gene.  
 CC  
 CC  
 SQ Sequence 1689 BP; 431 A; 416 C; 344 G; 497 T; 1 other;  
 XX  
 XX  
 XX Query Match 100.0%; Score 19; DB 18; Length 1689;  
 XX Best Local Similarity 100.0%; Pred. No. 6.2;  
 XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCAGTCTCTG 19  
 |||||  
 DB 285 CTGGGCCATCAGTCTCTG 267

RESULT 20  
 AAV07402/C  
 ID AAV07402 standard; DNA; 1689 BP.  
 XX  
 AC AAV07402;  
 XX  
 DT 28-SEP-1998 (first entry)  
 XX  
 DE Human C-C chemokine receptor 3 genomic DNA.

XX  
 XX C-C chemokine receptor 3; CRK-3; CCR3; Bos L2; human;  
 XX G protein-coupled receptor; leukocyte; antibody; antagonist;  
 XX inflammation; allergy; asthma; graft rejection; infection;  
 XX autoimmune disease; drug screening; therapy; ds.  
 OS  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers

FT CDS 181..1248  
 FT /\*tag= a  
 XX  
 PN MO9814480-A1.  
 PD 09-APR-1998.  
 XX  
 PF 24-SEP-1997; 97MO-US17103.  
 XX  
 PR 30-SEP-1996; 96US-0720565.  
 XX  
 XX (LEUK-) LEUKOSITE INC.  
 XX Mackay CR, Ponath PD;  
 XX WPI; 1998-286418/25.  
 XX P-PSDB; AAW51744.  
 DR  
 XX Antibodies to chemokine receptor-3 protein - useful for diagnosis  
 PT and treatment of inflammatory conditions, e.g. allergy, asthma,  
 PT autoimmune disease, graft rejection or cancer  
 XX  
 PS Example 2; Fig 1A-C; 185pp; English.  
 XX  
 CC This genomic DNA codes for novel human C-C chemokine receptor 3  
 CC (see AAW51744), also designated CKR-3, CCR3 or Eos 12, that binds and  
 CC mediates chemotaxis in response to chemokines such as eotaxin,  
 CC RANTES and MCP-3. The DNA was isolated from a human genomic phage  
 CC library using as probe a PCR fragment that had been generated from  
 CC eosinophil cDNA using primers (see AAV07405-12) based on known  
 CC chemokine receptor genes. A cDNA clone (see AAV07403) for CKR-3 is  
 CC also provided. Comparison of the sequences suggests that the  
 CC genomic DNA has an intron that separates the promoter and most  
 CC of the 5' untranslated region from the coding region. A consensus  
 CC of the genomic and cDNA sequences is provided (see AAV07404). The  
 CC invention relates to isolated and/or recombinant nucleic acids  
 CC encoding CKR-3, isolated or recombinant CKR-3 polypeptides,  
 CC recombinant nucleic acid constructs, host cells useful for  
 CC production of recombinant CKR-3 proteins, to antibodies reactive  
 CC with the receptors, and to methods of using these products to  
 CC identify ligands, antagonists and agonists of receptor function.  
 CC Inhibitors of CKR-3 can be used to treat: inflammatory or allergic  
 CC diseases and conditions, including respiratory allergic diseases  
 CC such as asthma, allergic rhinitis, hypersensitivity lung disease,  
 CC hypersensitivity pneumonitis, eosinophilic pneumonia (e.g.  
 CC Loeffler's syndrome, chronic eosinophilic pneumonia, interstitial  
 CC lung disease (ILD) e.g. idiopathic pulmonary fibrosis or ILD  
 CC associated with rheumatoid arthritis, systemic lupus erythematosus,  
 CC ankylosing spondylitis, systemic sclerosis, Sjogren's syndrome,  
 CC polymyositis or dermatomyositis), systemic anaphylaxis or  
 CC hypersensitivity responses, drug allergy, insect sting allergy,  
 CC inflammatory bowel disease, such as Crohn's disease and ulcerative  
 CC colitis, spondyloarthropathy, scleroderma, psoriasis, inflammatory  
 CC dermatosis such as dermatitis, eczema, atopic dermatitis,  
 CC allergic contact dermatitis, urticaria, vasculitis (e.g. necrotizing,  
 CC cutaneous and hypersensitivity vasculitis); eosinophilic myositis  
 CC and eosinophilic fasciitis; autoimmune diseases such as rheumatoid  
 CC arthritis, psoriatic arthritis, multiple sclerosis, systemic lupus  
 CC erythematosus, myasthenia gravis, juvenile onset diabetes,  
 CC glomerulonephritis, autoimmune thyroiditis and Behcet's disease;  
 CC graft rejection, including allograft rejection or graft-versus-host  
 CC disease; cancers with leukocyte infiltration of the skin or organs,  
 CC and also reperfusion injury, atherosclerosis, certain hematologic  
 CC malignancies, septic shock and endotoxic shock. Promoters of CKR-3  
 CC function can be used for treating: immunosuppression e.g. in AIDS  
 CC patients or individuals undergoing radiation therapy, chemotherapy,  
 CC therapy for autoimmune disease or other drug therapy, and  
 CC immunosuppression due congenital deficiency in receptor function or  
 CC other causes; and infectious diseases such as parasitic diseases,  
 CC including helminth infections, such as nematodes (round worms).  
 CC The agents can also be used for detection and diagnosis.  
 XX  
 .SQ Sequence 1689 BP; 431 A; 416 C; 344 G; 497 T; 1 other;

Query Match 100.0%; Score 19; DB 19; Length 1689;  
 Best Local Similarity 100.0%; Pred. No. 6.2;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGGCATCATGCTCTG 19  
 |||||  
 Db 285 CTGGGCATCATGCTCTG 267  
 RESULT 21  
 AAF21268/c  
 ID AAF21268 standard; DNA; 1689 BP.  
 XX  
 AC AAF21268;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Human low adenosine antisense oligonucleotide related sequence #2835.  
 XX  
 KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
 KW human; airway disorder; bronchoconstriction; lung inflammation;  
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cyostatic;  
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 KW cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200062736-A2.  
 XX  
 PD 26-OCT-2000.  
 XX  
 PF 24-MAR-2000; 2000MO-US08020.  
 XX  
 PR 06-APR-1999; 99US-0127958.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 PA (NYCE/) NYCE J W.  
 XX  
 PI NYce JW;  
 XX  
 DR WPI; 2000-679539/66.  
 XX  
 XX Low adenosine (A) content antisense oligonucleotides which do not  
 PT trigger adenosine receptors during metabolism, useful e.g. for treating  
 PT cancers and respiratory obstructions -  
 XX  
 PS Disclosure; Page 1183; 1592pp; English.  
 XX  
 CC The present invention describes low adenosine (A) content antisense  
 CC oligonucleotides and compositions (I) comprising them. In the antisense  
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 CC immunosuppressive, antiasthmatic, hypotensive and cyostatic activities.  
 CC The antisense oligonucleotides and (I) can be used to down-regulate the  
 CC expression and or activity of target polypeptides associated with  
 CC lung/respiratory disorders and malignancies, such as stimulating and  
 CC activating peptide factors and transmitters, transcription factors,  
 CC immunoglobulins and antibodies, antibody receptors, cytokines and  
 CC chemokines, endogenously produced specific and non-specific enzymes,  
 CC binding proteins, adhesion molecules and their receptors, cytokine and  
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
 CC nervous system (CNS) and peripheral nervous and non-nervous system  
 CC receptors, CNS and peripheral nervous and non-nervous system peptide  
 CC transmitters, defensins, growth factors, vasoactive peptides and  
 CC receptors, binding proteins and malignancy associated proteins. The  
 CC antisense oligonucleotides may be used in this way to treat disorders  
 CC including respiratory obstruction (especially pulmonary obstruction

CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)  
 CC and/or surfactant hypoproduction which are associated with a disease or  
 CC condition selected from pulmonary vasoconstriction, inflammation,  
 CC allergies, asthma, impaired respiration, respiratory distress syndrome  
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
 CC fragments and antisense oligonucleotides used in the exemplification of  
 CC the present invention.

XX Sequence 1689 BP; 430 A; 416 C; 345 G; 497 T; 1 other;  
 SQ

Query Match 100.0%; Score 19; DB 21; Length 1689;  
 Best Local Similarity 100.0%; Pred. No. 6.2;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCATGCTCTCTG 19  
 |||||  
 DB 285 CTGGGCCATCATGCTCTCTG 267

RESULT 22  
 AAA35146/C  
 ID AAA35146 standard; DNA; 1689 BP.  
 XX  
 AC AAA35146;  
 XX  
 DT 28-JUL-2000 (first entry)  
 XX  
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:20.  
 XX  
 KW Human: adenosine receptor; low adenosine antisense oligonucleotide;  
 KW phosphorothioate; impaired respiration; inflammation; allergy;  
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KW antiallergic; antiasthmatic; cyostatic; analgesic; impaired airway;  
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200009525-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 03-AUG-1999; 99WO-US17712.  
 XX  
 PR 03-AUG-1998; 98US-0095212.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI Nyce JW;  
 XX  
 DR MPI; 2000-205971/18.  
 XX  
 PT New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
 PT cancers -  
 XX  
 PS Disclosure; Page 1103; 1343pp; English.

CC The present invention describes a new composition comprising an  
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which  
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or  
 CC inflammation. The ON can have antiinflammatory, antiallergic,  
 CC antiasthmatic, cyostatic and analgesic activities. The compositions are  
 CC useful for the treatment of diseases associated with inflammation,  
 CC impaired airways, including lung disease and diseases whose secondary  
 CC effects afflict the lungs of a subject. They can be used for treating  
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,

CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic  
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
 CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,  
 CC carcinomas, and cancers which may metastasize to the lungs, including  
 CC breast and prostate cancer. The reduction of the adenosine content of  
 CC the ONs reduces side effects. The A-containing ONs break down with the  
 CC release of deoxyadenosine which activates adenosine receptors causing  
 CC bronchoconstriction and inflammation. AAA2313 to AAA35312 represent the  
 CC nucleotide sequences given in the sequence listing from the present  
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last  
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences  
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680  
 CC (AAA2323 to AAA3392) are specifically claimed ONs from the present  
 CC invention. N.B. Sequences given in the disclosure of the present  
 CC invention do not match up with their corresponding SEQ ID NO: sequences  
 CC given in the sequence listing.

XX Sequence 1689 BP; 430 A; 416 C; 345 G; 497 T; 1 other;  
 SQ

Query Match 100.0%; Score 19; DB 21; Length 1689;  
 Best Local Similarity 100.0%; Pred. No. 6.2;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGCCATCATGCTCTCTG 19  
 |||||  
 DB 285 CTGGGCCATCATGCTCTCTG 267

RESULT 23  
 ABL40462/C  
 ID ABL40462 standard; cDNA; 1689 BP.  
 XX  
 AC ABL40462;  
 XX  
 DT 10-JUN-2002 (first entry)  
 XX  
 DE Human C-C chemokine receptor 3 (CCR3) protein encoding cDNA.  
 XX  
 KW Mucosae-associated epithelial chemokine; MEC; C-C chemokine receptor;  
 KW CCR3; CCR10; anti-inflammatory; cyostatic; immunomodulator; anti-viral;  
 KW antibacterial; chemokine; human; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 181..1248  
 FT /\*tag= a  
 FT /product= "CCR3"  
 XX  
 PN WO200214532-A2.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 15-AUG-2001; 2001WO-US25734.  
 XX  
 PR 15-AUG-2000; 2000US-0638914.  
 XX  
 PA (MILL-) MILLBENNIUM PHARM INC.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 XX  
 PI Butcher EC, Kunkel EJ, Pan J, Soler-Ferran D;  
 XX  
 DR MPI; 2002-269204/31.  
 DR P-PSDB; ABB07733.  
 XX  
 PT Identifying modulators of mucosae-associated epithelial chemokine (MEC)  
 PT receptors 3 or 10 (CCR3/10), useful for treating inflammatory diseases,  
 PT comprising detecting formation of MEC-CCR3/10 complex or modulation of a  
 PT MEC-induced response -  
 XX  
 PS Example 2; Fig 4A-B; 92pp; English.  
 XX  
 CC The invention relates to identifying agents that inhibit or promote the

CC binding of a mammalian mucosae-associated epithelial chemokine (MEC) to  
CC a mammalian C-C chemokine receptor 3 (CCR3) or 10 (CCR10). The method  
CC involves: (a) detecting or measuring the formation of a complex between  
CC the MEC, and the CCR3 or CCR10; or (b) determining the ability of the  
CC test agent to inhibit or augment a MEC-induced response. An augmentation  
CC of complex formation, relative to a control, is indicative that the agent  
CC is a promoter. The method is useful for identifying modulators (e.g.  
CC inhibitors or promoters) of MEC-induced functions of CCR3 and/or CCR10.  
CC The inhibitors are useful for treating inflammatory diseases or  
CC conditions in a subject, e.g. oral inflammatory condition (e.g. Sjogren's  
CC syndrome or Behcet's syndrome), mastitis, chronic obstructive lung  
CC disease, asthma, inflammatory bowel disease (e.g. Crohn's disease,  
CC ulcerative colitis or coliac disease), IGA nephropathy or dermatitis  
CC herpeticiformis. The promoters are useful for treating cancers (e.g. solid  
CC tumours or cutaneous T cell lymphoma), neoplastic disease, retinopathy,  
CC macular degeneration, bacterial infections, tuberculous leprosy, viral  
CC infections, AIDS, neutropenias or bronchiectasis. The present sequence  
CC represents the human CCR3 protein encoding cDNA.

SQ Sequence 1689 BP; 430 A; 416 C; 345 G; 497 T; 1 other;

Query Match 100.0%; Score 19; DB 24; Length 1689;

Best Local Similarity 100.0%; Pred. No. 6.2; Mismatches 0; Gaps 0;

Matches 19; Conservative 0; Indels 0; Gaps 0;

1 CTGGGCGCATCATGTCCTCTG 19

285 CTGGGCGCATCATGTCCTCTG 267

#### RESULT 24

ABX13645/c

ID ABX13645 standard; cDNA; 1689 BP.

AC ABX13645;

DT 12-FEB-2003 (first entry)

XX Human cDNA encoding C-C chemokine receptor 3, CCR3.

KW Human; ss; gene; mucosae-associated epithelial chemokine; MEC;  
KW chromosome 5; C-C chemokine receptor; CCR3; CCR10; allergy;  
KW IGA antibody-secreting cell; inflammatory disease; Sjogren's syndrome;  
KW chronic obstructive lung disease; asthma; inflammatory bowel disease;  
KW IGA nephropathy; dermatitis herpeticiformis; ulcerative colitis;  
KW coeliac disease; autoimmune disease; arthritis; multiple sclerosis;  
KW systemic lupus erythematosus; myasthenia gravis; diabetes;  
KW autoimmune thyroiditis; graft rejection; atherosclerosis; cancer;  
KW infectious disease; immunosuppression.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 181..1248

FT /tag= a

FT /product= "CCR3"

XX US002137107-A1.

XX 26-SEP-2002.

XX 15-AUG-2001; 2001US-0931381.

XX 15-AUG-2000; 2000US-0638914.

XX (MILL-) MILLENNIUM PHARM INC.

XX Butcher EC, Kunkel EJ, Pan J, Soler-Ferran D;

XX WPI; 2003-102382/09.

XX P-PSDB; ABG72634.

PT Identifying agents that inhibit or promote binding of mammalian MEC to  
PT CCR3 or CCR10, useful for treating an inflammatory and autoimmune  
PT diseases, atherosclerosis, cancers, infectious diseases, diabetes,  
PT arthritis and asthma.

PS Disclosure; Fig 4; 46pp; English.

CC The invention relates to identifying an agent which inhibits or promotes  
CC the binding of mammalian MEC (mucosae-associated epithelial chemokine) to  
CC a mammalian C-C chemokine receptor 3 (CCR3) or CCR10 comprising:  
CC (a) combining an agent to be tested, a composition comprising a  
CC naturally occurring mammalian CCR3, CCR10 or MEC-binding variant, and a  
CC mammalian MEC; and (b) detecting or measuring the formation of a complex  
CC between the MEC and the CCR3, CCR10 or MEC-binding variant. The  
CC method alternatively comprises: (a) combining an agent to be tested, a  
CC cell expressing a protein comprising a naturally occurring mammalian  
CC CCR3, CCR10 or MEC-binding variant, and a mammalian MEC under conditions  
CC suitable for detecting an MEC-induced response; and (b) determining the  
CC ability of the test agent to inhibit or augment, where inhibition or  
CC augmentation of a MEC-induced response by the agent indicates that the  
CC agent is an inhibitor or promoter. Also included are: (1) an  
CC immunoglobulin or antigen-binding fragment which binds a naturally  
CC occurring mammalian CCR3 or CCR10, and inhibits the binding of a  
CC naturally occurring mammalian MEC to the receptor; and (2) modulating the  
CC activity of an IGA antibody-secreting cell in a subject, comprising  
CC administering MEC or an agent that promotes or inhibits the binding of  
CC MEC to CCR3 and/or CCR10. The methods and compositions of the present  
CC invention are useful for treating an inflammatory disease or allergic  
CC condition such as oral inflammatory condition (e.g. Sjogren's or  
CC Behcet's syndrome), mastitis, chronic obstructive lung disease, asthma,  
CC inflammatory bowel disease, IGA nephropathy, dermatitis herpeticiformis,  
CC ulcerative colitis and coeliac disease. They can also be used in  
CC autoimmune diseases such as arthritis, multiple sclerosis,  
CC systemic lupus erythematosus, myasthenia gravis, diabetes and autoimmune  
CC thyroiditis, graft rejection, atherosclerosis, cancers, infectious  
CC diseases and immunosuppression (many other diseases and disorders are  
CC listed in the specification). The gene for MEC is located on chromosome  
CC 5. The present sequence is the Human cDNA encoding C-C chemokine  
CC receptor 3, CCR3.

XX Sequence 1689 BP; 430 A; 416 C; 345 G; 497 T; 1 other;

Query Match 100.0%; Score 19; DB 25; Length 1689;

Best Local Similarity 100.0%; Pred. No. 6.2; Mismatches 0; Gaps 0;

Matches 19; Conservative 0; Indels 0; Gaps 0;

1 CTGGGCGCATCATGTCCTCTG 19

285 CTGGGCGCATCATGTCCTCTG 267

#### RESULT 25

ABL67066/c

ID ABL67066 standard; DNA; 1717 BP.

AC ABL67066;

DT 15-MAY-2002 (first entry)

XX Thyroid cancer related gene sequence SEQ ID NO:5403.

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytosaratic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236103P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237284P.  
 PR 02-OCT-2000; 2000US-237285P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 XX  
 PA (AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Sopet DR, Weaver Z;  
 XX  
 XX MPI; 2002-188264/24.  
 DR  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set  
 XX  
 PS Claim 1; SEQ ID 5403; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine

CC carcinoma, papillary carcinoma and Wilm's tumour.  
 XX  
 XX Sequence 1717 BP; 434 A; 428 C; 351 G; 504 T; 0 other;  
 Query Match 100.0%; Score 19; DB 24; Length 1717;  
 Best Local Similarity 100.0%; Pred. No. 6.3;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 CTGGGCGCATCAGTCTCTG 19  
 Db 309 CTGGGCGCATCAGTCTCTG 291  
 RESULT 26  
 AAD25221/c  
 ID AAD25221 standard; DNA; 1717 BP.  
 XX  
 AC AAD25221;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Human chemokine (C-C motif) receptor 3 (CCR3) gene #1.  
 XX  
 KW Human; chemokine (C-C motif) receptor 3; CCR3 gene; haplotyping;  
 KW genotyping; type IV hypersensitivity reaction; HIV-1; gene therapy;  
 KW human immunodeficiency virus 1; single nucleotide polymorphism; SNP;  
 KW chromosome 3p21.3; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH variation replace (92, T)  
 FT /tag= a  
 FT /standard name= "Single nucleotide polymorphism (SNP)"  
 FT variation replace (197, A)  
 FT /tag= b  
 FT /standard name= "Single nucleotide polymorphism (SNP)"  
 FT CDS /tag= c  
 FT /note= "This region corresponds to exon 1"  
 FT variation replace (255, C)  
 FT /tag= d  
 FT /standard name= "Single nucleotide polymorphism (SNP)"  
 FT variation replace (1256, C)  
 FT /tag= e  
 FT /standard name= "Single nucleotide polymorphism (SNP)"  
 XX  
 PD WO200187908-A2.  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001WO-US16278.  
 XX  
 PF 18-MAY-2000; 2000US-205191P.  
 XX  
 PA (GENA-) GENAISSANCE PHARM INC.  
 XX  
 PI Choi JY, Kazemi A, Koshy B;  
 PI WPI; 2002-055681/07.  
 DR P-PSDB; AAE15320.  
 XX  
 PT Isolated polymorphic variants of chemokine (C-C motif) receptor 3  
 PT (CCR3) gene useful for studying function of CCR3, expressing the CCR3  
 PT protein and to screen drugs to treat CCR3 activity-related diseases -  
 XX  
 PS Example 1; Fig 1; 53pp; English.  
 XX  
 CC The invention relates to genetic variants of human chemokine (C-C motif)  
 CC receptor 3 (CCR3) gene. The invention also relates to compositions and  
 CC methods for haplotyping and/or genotyping the CCR3 gene in an individual.  
 CC Polynucleotides of the invention are useful for studying the expression





XX Claim 7; Page 48-50; 65pp; English.

PS This sequence comprises a full-length cDNA coding for novel human

CC chemokine receptor 88-2B (AAW27124), a G protein coupled receptor that

CC is involved in leukocyte trafficking. The 88-2B cDNA was obtained

CC from a macrophage cDNA library using 88-2B-specific primers. A

CC full-length clone (see AAT89161) for chemokine receptor 88C (AAW27123)

CC was also obtained. 88C and 88-2B cDNAs can be used to produce

CC recombinant polypeptides in transformed host cells for use in the

CC treatment of e.g. atherosclerosis, rheumatoid arthritis, tumors,

CC asthma, viral infection, AIDS and inflammatory conditions. Nucleic

CC acid fragments can be used to isolate genomic sequences, to detect

CC alleles of the gene (for diagnosis or in gene therapy), to alter

CC receptor genetics to facilitate identification of modulators and to

CC produce knockout animals, and (antisense forms) to alter/study the

CC genetics and expression of the receptor.

SQ Sequence 1915 BP; 488 A; 470 C; 373 G; 584 T; 0 other;

Query Match 100.0%; Score 19; DB 18; Length 1915;

Best Local Similarity 100.0%; Pred. NO. 6.4; Mismatches 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGGGCCATCAGTGTCTCTG 19

|||||

Db 466 CTGGGGCCATCAGTGTCTCTG 448

RESULT 29

ABT04010/c

ID ABT04010 standard; DNA; 3426 BP.

XX

XX ABT04010;

AC

XX 25-SEP-2002 (first entry)

DT

XX

XX Human ovary specific coding sequence SEQ ID NO: 29.

DE

XX

XX Human; ovary; ovarian cancer; ovarian disease; gene therapy; gene;

KM cytoskeletal; ds.

KX

XX

OS Homo sapiens.

XX

XX WO200240720-A2.

PN

XX

PD 23-MAY-2002.

PT

XX

PF 20-NOV-2001; 2001WO-US45010.

PS

XX

PR 20-NOV-2000; 2000US-249997P.

XX

XX (DIAD-) DIADEXUS INC.

PA

XX

PI Salceda S, Macina RA, Recipon H, Cafferty R, Sun Y, Liu C;

XX

DR WPI; 2002-547586/58.

XX

XX New ovary polypeptides useful for detecting, diagnosing, monitoring,

PT treating, staging and imaging cancers in humans having cancer and

PT non-cancerous ovary disease

XX

PS Claim 1; Page 162-164; 296pp; English.

XX

XX The present invention provides human proteins and coding sequences

CC specifically found in ovary cells. These can be used in the diagnosis and

CC treatment of ovarian diseases, including cancer. The present sequence is

CC a coding sequence of the invention.

XX

SQ Sequence 3426 BP; 1107 A; 828 C; 714 G; 777 T; 0 other;

Query Match 100.0%; Score 19; DB 24; Length 3426;

Best Local Similarity 100.0%; Pred. No. 6.9;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGGGCCATCAGTGTCTCTG 19

|||||

Db 234 CTGGGGCCATCAGTGTCTCTG 216

RESULT 30

AAF21269/c

ID AAF21269 standard; DNA; 3958 BP.

XX

XX AAF21269;

AC

XX

XX 14-MAR-2001 (first entry)

DT

XX

XX Human low adenosine antisense oligonucleotide related sequence #2836.

DE

XX

XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;

KM human; airway disorder; bronchoconstriction; lung inflammation;

KM surfactant depletion; respiratory; bronchodilator; antiinflammatory;

KM immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;

KM respiratory obstruction; pulmonary obstruction; impeded respiration;

KM surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;

KM respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;

KM pulmonary hypertension; emphysema; pulmonary transplantation rejection;

KM chronic obstructive pulmonary disease; pulmonary infection; bronchitis;

KM cancer; ss.

XX

XX

OS Homo sapiens.

XX

XX WO200062736-A2.

PN

XX

PD 26-OCT-2000.

PT

XX

PF 24-MAR-2000; 2000WO-US08020.

PS

XX

PR 06-APR-1999; 99US-0127958.

XX

XX (IYEC-) UNIV EAST CAROLINA.

PA

XX

PI (NYCE/) NYCE J W.

XX

PI Nyce JW;

XX

XX WPI; 2000-679539/66.

DR

XX

XX Low adenosine (A) content antisense oligonucleotides which do not

PT trigger adenosine receptors during metabolism, useful e.g. for treating

PT cancers and respiratory obstructions

XX

XX

PS Disclosure; Page 1183-1184; 1592pp; English.

XX

XX The present invention describes low adenosine (A) content antisense

CC oligonucleotides and compositions (I) comprising them. In the antisense

CC oligonucleotides the A is replaced by a 'universal' or alternative base.

CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,

CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.

CC The antisense oligonucleotides and (I) can be used to down-regulate the

CC expression and or activity of target polypeptides associated with

CC lung/respiratory disorders and malignancies, such as stimulating and

CC activating peptide factors and transmitters, transcription factors,

CC immunoglobulins and antibodies, antibody receptors, cytokines and

CC chemokines, endogenously produced specific and non-specific enzymes,

CC binding proteins, adhesion molecules and their receptors, cytokine and

CC chemokine receptors, adenosine receptors, bradykinin receptors, central

CC nervous system (CNS) and peripheral nervous and non-nervous system

CC receptors, CNS and peripheral nervous and non-nervous system peptide

CC transmitters, defensins, growth factors, vasoactive peptides and

CC receptors, binding proteins and malignancy associated proteins. The

CC antisense oligonucleotides may be used in this way to treat disorders

CC including respiratory obstruction (especially pulmonary obstruction

CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)

CC and/or surfactant hypoproduction which are associated with a disease or

CC condition selected from pulmonary vasoconstriction, inflammation,

CC allergies, asthma, impeded respiration, respiratory distress syndrome  
 CC (BDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
 CC fragments and antisense oligonucleotides used in the exemplification of  
 CC the present invention.

XX  
 SQ Sequence 3958 BP; 939 A; 1025 C; 855 G; 1138 T; 1 other;

Query Match 100.0%; Score 19; DB 21; Length 3958;  
 Best Local Similarity 100.0%; Pred. No. 7.1;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCGCATCAGTCTCTG 19  
 |||||  
 DB 2554 CTGGGCGCATCAGTCTCTG 2536

RESULT 31  
 AAA35147/c  
 ID AAA35147 standard; DNA; 3958 BP.  
 XX  
 AC AAA35147;  
 XX  
 DT 28-JUL-2000 (first entry)  
 XX  
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:21.  
 XX  
 XX Human; adenosine receptor; low adenosine antisense oligonucleotide;  
 KW phosphorothioate; impaired respiration; inflammation; allergy;  
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KW antiallergic; antiasthmatic; cytosolic; analgesic; impaired airway;  
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200009525-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 03-AUG-1999; 99WO-US17712.  
 XX  
 PR 03-AUG-1998; 98US-0095212.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI NYCE JW;  
 XX  
 DR WPI; 2000-205971/18.  
 XX  
 PT New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
 PT cancers -  
 XX  
 XX Disclosure; Page 1103-1104; 1343BP; English.

CC The present invention describes a new composition comprising an  
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which  
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or  
 CC inflammation. The ON can have antiinflammatory, antiallergic,  
 CC antiasthmatic, cytostatic and analgesic activities. The compositions are  
 CC useful for the treatment of diseases associated with inflammation,  
 CC impaired airways, including lung disease and diseases whose secondary  
 CC effects afflict the lungs of a subject. They can be used for treating  
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,  
 CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic  
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
 CC pulmonary disease (COPD), and cancers such as leukaemia, lymphomas,

CC carcinomas, and cancers which may metastasise to the lungs, including  
 CC breast and prostate cancer. The reduction of the adenosine content of  
 CC the ONs reduces side effects. The A-containing ONs break down with the  
 CC release of deoxyadenosine which activates adenosine receptors causing  
 CC bronchoconstriction and inflammation. AAA23313 to AAA35312 represent the  
 CC nucleotide sequences given in the sequence listing from the present  
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last  
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences  
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680  
 CC (AAA3323 to AAA3392) are specifically claimed ONs from the present  
 CC invention. N.B. Sequences given in the disclosure of the present  
 CC invention do not match up with their corresponding SEQ ID NO: sequences  
 CC given in the sequence listing.

XX  
 SQ Sequence 3958 BP; 939 A; 1025 C; 855 G; 1138 T; 1 other;

Query Match 100.0%; Score 19; DB 21; Length 3958;  
 Best Local Similarity 100.0%; Pred. No. 7.1;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGGCGCATCAGTCTCTG 19  
 |||||  
 DB 2554 CTGGGCGCATCAGTCTCTG 2536

RESULT 32  
 AAT93601/c  
 ID AAT93601 standard; cDNA; 5099 BP.  
 XX  
 AC AAT93601;  
 XX  
 DT 07-MAY-1998 (first entry)  
 XX  
 DE Human eosinophil ectaxin receptor CC CKR3 encoding cDNA.  
 XX  
 KW Eosinophil ectaxin receptor; CC CKR3; human; treatment; dermatitis;  
 KW atopic condition; allergic rhinitis; conjunctivitis; bronchial asthma;  
 KW beta-chemokine receptor; viral infection; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9741154-A1.  
 XX  
 PD 06-NOV-1997.  
 XX  
 PF 24-APR-1997; 97WO-US06568.  
 XX  
 PR 17-JAN-1997; 97GB-0000894.  
 PR 26-APR-1996; 96US-0016158.  
 PR 26-APR-1996; 96US-0017113.  
 XX  
 PA (MERI ) MERCK & CO INC.  
 XX  
 PI Daugherty BL, Demartino JA, Siciliano SJ, Springer MS;  
 XX  
 DR WPI; 1997-549685/50.  
 XX  
 PT New isolated human eosinophil ectaxin receptor - used to develop  
 PT products for treating and preventing atopic conditions e.g. allergic  
 PT rhinitis, dermatitis, conjunctivitis and bronchial asthma

PS Claims 12, 13, 14; Pages 16-20; 51pp; English.  
XX This cDNA encodes a human eosinophil ectoxin receptor. This 5099 base  
CC pair sequence comprises a 1065 base pair open reading frame encoding a  
CC 355 amino acid eosinophil ectoxin receptor protein, flanked by a 5'  
CC genomic DNA sequence and a 3' terminator region. This novel eosinophil  
CC ectoxin receptor is a human beta-chemokine receptor designated CC CCR3.  
CC Agents which bind to this eosinophil ectoxin receptor can be used for  
CC the treatment and prevention of atopic conditions such as allergic  
CC rhinitis, dermatitis, conjunctivitis and bronchial asthma. Agents which  
CC block this eosinophil ectoxin receptor can be used to prevent viral  
CC infection in healthy individuals and slow or halt viral progression  
CC in infected patients.  
XX  
SQ Sequence 5099 BP; 1388 A; 1171 C; 1013 G; 1527 T; 0 other;  
XX  
Query Match 100.0%; Score 19; DB 18; Length 5099;  
Best Local Similarity 100.0%; Pred. No. 7.3;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CTGGGCGCATCAGTGTCTG 19  
Db 3691 CTGGGCGCATCAGTGTCTG 3673  
XX  
RESULT 33  
ABZ68879/c  
ID ABZ68879 standard; cDNA; 5791 BP.  
XX  
AC ABZ68879;  
XX  
DT 28-MAY-2003 (first entry)  
XX  
DE Nucleotide sequence of human chemokine receptor CCR3.  
XX  
KW Human; chemokine receptor; CCR3; viral infection; surface protein;  
KW respiratory virus infection; respiratory syncytial virus infection;  
KW RSV infection; bronchiolitis; bronchitis; pneumonia; asthma; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 4015..5082  
FT /\*tag= a  
FT /product= "CCR3"  
XX  
PN WO2003014153-A2.  
XX  
PD 20-FEB-2003.  
XX  
PF 12-AUG-2002; 2002WO-CA01248.  
XX  
PR 10-AUG-2001; 2001US-311088P.  
XX  
PA (TOP1-) TOPIGEN PHARM INC.  
XX  
PI Renzl P, Zemzouni K, Lamkhoulou B;  
XX  
DR WPI; 2003-247991/25.  
DR P-PSDB; ABP97726.  
XX  
PT Modulating viral infection of a cell, for treating or preventing  
PT respiratory virus infections, bronchitis, pneumonia or asthma, by  
PT modulating a binding interaction between a cell chemokine-receptor and  
PT a surface protein of the virus  
XX  
XX Disclosure; Page 84-88; 120pp; English.  
XX  
XX The present sequence encodes human chemokine receptor CCR3. The  
CC specification describes a method for modulating viral infection of  
CC a cell. The method comprises modulating a binding interaction between  
CC a cell chemokine-receptor and a surface protein of the virus. The  
CC proviso is that the cell chemokine-receptor is not CX3CR1 and that the

CC virus is not HIV. The method is useful for treating or preventing  
CC respiratory virus infection in vertebrates, more particularly  
CC respiratory syncytial virus (RSV) infections, and related diseases,  
CC e.g. bronchiolitis, bronchitis, pneumonia or asthma.  
XX  
SQ Sequence 5791 BP; 1624 A; 1285 C; 1160 G; 1722 T; 0 other;  
XX  
Query Match 100.0%; Score 19; DB 25; Length 5791;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CTGGGCGCATCAGTGTCTG 19  
Db 4119 CTGGGCGCATCAGTGTCTG 4101  
XX  
RESULT 34  
ABA63970/c  
ID ABA63970 standard; DNA; 549 BP.  
XX  
AC ABA63970;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human foetal liver single exon nucleic acid probe #12275.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human foetal liver -  
XX  
XX Claim 1; SEQ ID NO 12275; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC foetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 549 BP; 127 A; 148 C; 178 G; 96 T; 0 other;  
XX  
Query Match 91.6%; Score 17.4; DB 22; Length 549;  
Best Local Similarity 94.7%; Pred. No. 35;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 CTGGGCGCATCAGTGTCTG 19  
XX



XX 30-JAN-2001; 2001WO-US00668.  
XX  
PF 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR WPI; 2001-488900/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
PS Example 4; SEQ ID NO: 12736; 658bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukemia and myeloma. The present sequence is one of  
CC the probes of the invention.  
XX  
SQ Sequence 549 BP; 127 A; 148 C; 178 G; 96 T; 0 other;  
XX  
Query Match 91.6%; Score 17.4; DB 22; Length 549;  
Best Local Similarity 94.7%; Pred. No. 35;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CTGGGCCCATGAGTGTCTTG 19  
DB 101 CTGGGCCCATGAGTGTCTTG 83  
XX  
RESULT 38  
AA118973/c  
ID AA118973 standard; DNA; 549 BP.  
XX  
AC AA118973;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Probe #8906 for gene expression analysis in human cervical cell sample.  
XX  
KW Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00670.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX

DR WPI; 2001-488901/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
PS Claim 25; SEQ ID No 8906; 487bp; English.  
XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENPs). The present sequence is one such probe. The SENPs are derived  
CC from human HeLa cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging  
CC of diseases of the cervix, notably cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 549 BP; 127 A; 148 C; 178 G; 96 T; 0 other;  
XX  
Query Match 91.6%; Score 17.4; DB 22; Length 549;  
Best Local Similarity 94.7%; Pred. No. 35;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CTGGGCCCATGAGTGTCTTG 19  
DB 101 CTGGGCCCATGAGTGTCTTG 83  
XX  
RESULT 39  
AA144108/c  
ID AA144108 standard; DNA; 549 BP.  
XX  
AC AA144108;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Probe #12794 used to measure gene expression in human placenta sample.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488897/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
PS Claim 25; SEQ ID No 12794; 654bp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENPs).  
CC The present sequence is one such probe. The probes are useful for  
CC producing a microarray for predicting, measuring and displaying gene  
CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders.

XX Sequence 549 BP; 127 A; 148 C; 178 G; 96 T; 0 other;

Query Match 91.6%; Score 17.4; DB 22; Length 549;

Best Local Similarity 94.7%; Pred. No. 35;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CTGGGCCATCAGTGTCTCTG 19

101 CTGGGCCATCAGTGTCTCTG 83

RESULT 40

ABS37802/C

ID ABS37802 standard; DNA; 549 BP.

XX ABS37802;

XX 25-FEB-2003 (first entry)

XX Human liver single exon probe, SEQ ID No 12792.

XX Human; single exon nucleic acid probe; liver; cirrhosis;

KM hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;

KM coronary heart disease; ss.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analysing gene expression in human adult liver -

XX Claim 1; SEQ ID No 12792; 658bp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (1) for

XX measuring human gene expression in a sample derived from human adult

XX liver, comprising one of 13109 defined nucleotide sequences given in the

XX specification (or complements/fragments). The probe hybridises at high

XX stringency to a nucleic acid molecule expressed in the human adult

XX liver. (1) may be used for predicting, measuring and displaying gene

XX expression in samples derived from human adult liver. The genes

XX identified may be involved in genetic liver diseases such as cirrhosis,

XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which

XX is associated with coronary heart disease. ABS25011-ABS51005 represent

XX human liver single exon nucleic acid probes of the invention.

XX Note: The sequence information for this patent does not appear in the

XX printed specification but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 549 BP; 127 A; 148 C; 178 G; 96 T; 0 other;

XX Query Match 91.6%; Score 17.4; DB 22; Length 549;

XX Best Local Similarity 94.7%; Pred. No. 35;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CTGGGCCATCAGTGTCTCTG 19

101 CTGGGCCATCAGTGTCTCTG 83

Search completed: January 13, 2004, 18:02:40  
Job time : 124.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 13, 2004, 15:53:18 : Search time 1192.5 Seconds  
(without alignments)  
387.241 Million cell updates/sec

Title: US-09-719-737-9

Perfect score: 19  
Sequence: 1 99gtcgcgcagcgagatggt 19

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_estba: \*  
3: em\_estba: \*  
4: em\_estba: \*  
5: em\_estba: \*  
6: em\_estba: \*  
7: em\_estba: \*  
8: em\_estba: \*  
9: em\_estba: \*  
10: em\_estba: \*  
11: em\_estba: \*  
12: em\_estba: \*  
13: em\_estba: \*  
14: em\_estba: \*  
15: em\_estba: \*  
16: em\_estba: \*  
17: em\_estba: \*  
18: em\_estba: \*  
19: em\_estba: \*  
20: em\_estba: \*  
21: em\_estba: \*  
22: em\_estba: \*  
23: em\_estba: \*  
24: em\_estba: \*  
25: em\_estba: \*  
26: em\_estba: \*  
27: em\_estba: \*  
28: em\_estba: \*  
29: em\_estba: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	299	10	BF884753 PM3-ET020
2	19	100.0	354	12	BM146496 TCAP1E49
3	19	100.0	357	10	BF986303 OVA-GN014
4	19	100.0	478	13	BU431613 UI-HF-BNO

Result No.	Score	Query Match	Length	ID	Description
5	19	100.0	785	14	CB961169 AGENCOURT
6	19	100.0	1031	13	BX359469 BX359469
7	17.4	91.6	209	14	H31083 EST104774 R
8	17.4	91.6	266	10	BB306760 BB306760
9	17.4	91.6	318	13	BQ311140 BQ311140
10	17.4	91.6	355	10	BE102744 UI-R-BT1-
11	17.4	91.6	389	9	AA892154 EST195957
12	17.4	91.6	463	9	CB735745 AMGNNUC:N
13	17.4	91.6	573	14	CB614936 AMGNNUC:N
14	17.4	91.6	580	28	AZ952841 2M0217N18
15	17.4	91.6	1077	12	B1114889 602861478
16	17.4	91.6	1164	12	BQ059970 AGENCOURT
17	17.4	91.6	1562	10	BE887786 601511359
18	17.4	91.6	3874	11	AK031524 Mue muscu
19	17.4	89.5	993	29	B2554088 pac61-60
20	16.4	86.3	206	12	BJ239259 BJ239259
21	16.4	86.3	214	9	AW612888 h317901.x
22	16.4	86.3	240	9	AJ460522 AJ460522
23	16.4	86.3	253	12	BJ293359 BJ293359
24	16.4	86.3	284	12	BJ299950 BJ299950
25	16.4	86.3	282	12	BJ299963 BJ299963
26	16.4	86.3	292	14	CA644225 wreln.pko
27	16.4	86.3	324	12	B1013302 PM4-ET015
28	16.4	86.3	336	13	BU974641 HB28024r
29	16.4	86.3	350	14	R25769 R25769
30	16.4	86.3	355	9	AW769047 h157h11.x
31	16.4	86.3	362	9	A1473980 tmd4609.x
32	16.4	86.3	369	10	BF112016 BF112016
33	16.4	86.3	380	13	BQ189288 UI-B-EJ1-
34	16.4	86.3	382	9	AA662148 h120B08.x
35	16.4	86.3	382	9	AA613554 nq22e02.8
36	16.4	86.3	394	10	BE411862 ISCO10.A
37	16.4	86.3	400	12	BM443000 EBR002 SQ
38	16.4	86.3	402	14	CB871117 HC02E21Y
39	16.4	86.3	403	28	BH306574 CH230-192
40	16.4	86.3	406	9	A1492104 t907b01.x
41	16.4	86.3	413	12	B1499969 l612h06.x
42	16.4	86.3	417	13	BX103590 BX103590
43	16.4	86.3	419	10	BE411823 ISCO09.F0
44	16.4	86.3	419	14	CB867914 HC02E21w
45	16.4	86.3	420	14	CA614992 w1.PK149

## ALIGNMENTS

RESULT 1  
LOCUS BF884753  
DEFINITION PM3-ET0207-271200-003-e12 ET0207 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF884753.1 GI:12275402  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
REFERENCE  
AUTHORS  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL MEDLINE  
PROC. NATL. ACAD. SCI. U.S.A. 97 (7), 3491-3496 (2000)  
PUBMED 10737800  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

FEATURES	source	1. .299 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /dev_stage="Adult" /clone_lib="ET0207" /note="Organ: lung tumor; Vector: puc18; Site 1: Smal; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent applications No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription cDNA tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT	50 a 86 c 79 g 84 t	
ORIGIN		
Query Match	100.0%; Score 19; DB 10; Length 299;	
Best Local Similarity	100.0%; Pred. No. 3.8e+02;	
Matches	19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 GGGTCTGCAGCGGATGCT 19 	
Db	270 GGGTCTGCAGCGGATGCT 288	
RESULT 2		
BM146496/c	354 bp mRNA linear EST 30-NOV-2001	
LOCUS	TCAP4963 Pediatric acute myelogenous leukemia cell (PAB M1)	
DEFINITION	Baylor-HGSC project=TCAP Homo sapiens cDNA clone TCAP4963, mRNA sequence.	
ACCESSION	BM146496	
VERSION	BM146496.1 GI:17165472	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 354)	
AUTHORS	Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr., Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.	
TITLE	Pediatric Leukemia cDNA Sequencing Project (2001)	
JOURNAL	Unpublished	
COMMENT	Contact: Dr. Judith F. Margolin Texas Children's Cancer Center and Human Genome Sequencing Center at Baylor College of Medicine 1102 Bates, MC3-3320 Houston, TX 77030, USA Tel: 832-824-4536 Fax: 832-825-4038 Email: clones@tcxccc.org Seq primer: M13 primer.	
FEATURES	source	1. .354 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_lib="TCAP4963" /sex="male" /tissue_type="leukopheresis" /cell_type="myeloid cell" /dev_stage="pediatric 6 years"

```

/lab host="DH10B"
/clone.lib="Pediatric acute myelogenous leukemia cell (PAB
M1) Baylor-HGSC project-TCNA"
/nc=vector: lambda PSB; Site_1: BamH1; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo (dT) primer (5'GGAGGACTCTGACCGCCGAGGAG(17)VN
3'; V=A,C,G; N=A,C,G,T) and then dg tailed. Second strand
was primed with a BamH1-dC primer
[5'AGAGAGCTCGATCCGCCGCCGCAATATATAT(3)3'].
Double-stranded cDNA was then digested with BamH1 and XhoI
and directionally cloned into the BamH1 and SalI sites of
lambda psb vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P. Westover A. Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap tripper.,
DNA Res 4: 1, 61-6, Feb 28, 1997")
BASE COUNT
77 a 113 c 106 g. 58 t
ORIGIN
Query Match 100.0%; Score 19; DB 12; Length 354;
Best Local Similarity 100.0%; Prid. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 0; Gaps 0;
QY 1 GGGCTGCGAGCGGATGTT 19
|||||
db 320 GGGCTGCGAGCGGATGTT 302
RESULT 3
BP986303/ 357 bp mRNA linear EST 23-JAN-2001
LOCUS BP986303
DEFINITION QV4-GN0144-031000-446-b10 GN0144 Homo sapiens cDNA, mRNA sequence.
ACCESSION BP986303
VERSION BP986303.1 GI:12392610
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 357)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,D.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
2020263
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Frudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPSP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&ct2=QV4-GN0144-
031000-446-b10&ct3=2000-10-03&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 5
High quality sequence stop: 57.
Location/Qualifiers
1. 357
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```



/dev\_stage="Adult"  
/clone\_id="GN0144"  
/note="Organ: placenta normal; Vector: puc18; Site\_1: SmaI  
; Site\_2: SmaI; A mini-library was made by cloning  
products derived from OREBRES PCR (U.S. Letters Patent  
application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the pUC 18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."

# BASE COUNT

77 a 110 c 116 g 54 t

# ORIGIN

Query Match 100.0%; Score 19; DB 10; Length 357;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
DB 316 GGGTCTGCAGCGGATGCT 298

# RESULT 4

BU431613/c 478 bp mRNA linear EST 09-SEP-2002  
LOCUS UI-HF-BN0-af-f-04-0-UI.r1 NIH\_MGC\_50 Homo sapiens cDNA clone  
DEFINITION IMAGE:3067902 5', mRNA sequence.

ACCESSION BU431613  
VERSION BU431613.1 GI:22770100  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS NIH-MGC  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bdrrp/image/image.html  
Seq primer: M13 Forward.

# FEATURES

## source

1..478  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3067902"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/cell\_line="MGC85"  
/lab\_host="DH10B (LTI)"  
/clone\_id="NIH\_MGC\_50"  
/note="Vector: pT73-Pac; Site\_1: NotI; Site\_2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(3.5-4.4kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Fatima Bonaldo, Ph.D. and M. Benito Soares, Ph.D."

# BASE COUNT

97 a 158 c 148 g 75 t

# ORIGIN

Query Match 100.0%; Score 19; DB 13; Length 478;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||||

DB 303 GGGTCTGCAGCGGATGCT 285

# RESULT 5

CB961169/c 785 bp mRNA linear EST 29-APR-2003  
LOCUS AGENCOURT\_13762074 NIH\_MGC\_147 Homo sapiens cDNA clone  
DEFINITION IMAGE:303444148 5', mRNA sequence.

ACCESSION CB961169  
VERSION CB961169.1 GI:30217286  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS NIH-MGC  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Dr. Stefan Hanson  
CDNA Library Preparation: Michael U. Brownstein (NHGRI) with help  
and advice from Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDAM379 row: p column: 13  
High quality sequence stop: 651.

# FEATURES

## source

1..785  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30344148"  
/tissue\_type="Human Placenta"  
/lab\_host="DH10B TONa"  
/clone\_id="NIH\_MGC\_147"  
/note="Organ: placenta; Vector: pBluescriptPR; Site\_1:  
all-XhoI; Site\_2: BamH; Oligo-dT primed using primer  
5'-TTTTTTTTTTTTTNN-3', size-selected for average  
insert size 2.3 kb and normalized to ROP 5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIH/NHGRI, National Institutes of Health). Note: This is  
a NIH MGC library."

# BASE COUNT

154 a 257 c 234 g 140 t

Query Match 100.0%; Score 19; DB 14; Length 785;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
DB 341 GGGTCTGCAGCGGATGCT 323

# RESULT 6

BX359469/c 1031 bp mRNA linear EST 05-MAY-2003  
LOCUS BX359469 Homo sapiens PLACENTA\_COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION BX359469 IMAGE:30382361  
ACCESSION BX359469  
VERSION BX359469.1 GI:30382361  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1031)  
 AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 10421.f. For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CSOD1057DD01Q1P1c1cluster=10421.f. Contact :  
 Feng Liang Email: fliang@life-tech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CSOD1057DD01Q1P1.

FEATURES  
 source  
 1..1031  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSOD1057RH02"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_1b="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and BcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 191 a 336 c 293 g 204 t  
 ORIGIN

Query Match 100.0%; Score 19; DB 13; Length 1031;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
 |||||  
 Db 202 GGGTCTGCAGCGGATGCT 184

RESULT 7  
 H31083/c 209 bp mRNA linear EST 02-APR-1998  
 LOCUS EST004774 Rat PC-12 cells, untreated Rattus sp. cDNA clone RCCAB19  
 DEFINITION 5' end, mRNA sequence.  
 H31083  
 H31083.1 GI:976505

ACCESSION  
 VERSION H31083.1  
 KEYWORDS EST.  
 SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 209)  
 AUTHORS Lee, N.H., Weinstein, K.G., Kirkness, E.F., Earle-Hughes, J.A., Fulmer,  
 R.A., Marmarosa, S., Glodet, A., Gocayne, J.D., Adams, M.D., Kerlavage  
 , A.R., Fraser, C.M., and Venter, J.C.  
 TITLE Comparative expressed-tag analysis of differential gene  
 expression profiles in PC-12 cells before and after nerve growth  
 factor treatment  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)  
 MEDLINE 95366786  
 PUBMED 7667285

COMMENT Other ESTs: EST104773  
 Contact: Lee, NH  
 The Institute for Genomic Research  
 9710, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 For clone availability please contact the TIGR Database  
 (cdinfo@tigr.org)  
 Seq primer: M13 Reverse.

FEATURES  
 source  
 Location/Qualifiers  
 1..209  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="ATCC (inhost):2000079"  
 /db\_xref="taxon:10118"  
 /clone="RCCAB19"  
 /clone\_1b="Rat PC-12 cells, untreated"  
 /note="Vector: pluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI; poly(A) + RNA was purified from untreated PC12 cells  
 cultured for 9 days. cDNA was constructed using an  
 oligo-dT primer and directionally cloned using the Lambda  
 Zap II Vector Kit by Stratagene"

BASE COUNT 48 a 61 c 51 g 49 t  
 ORIGIN

Query Match 91.6%; Score 17.4; DB 14; Length 209;  
 Best Local Similarity 94.7%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
 |||||  
 Db 194 GGGTCTGCAGCGGATGCT 176

RESULT 8  
 BB306760/c 266 bp mRNA linear EST 10-JUL-2000  
 LOCUS BB306760 RIKEN full-length enriched, adult male corpora  
 quadrigemina Mus musculus cDNA clone B230209H01.3 similar to  
 U79774 Mus musculus NNP-1 var (NNP-1) mRNA, mRNA sequence.  
 BB306760  
 BB306760.1 GI:9007465

ACCESSION  
 VERSION BB306760  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 266)  
 AUTHORS Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci,  
 P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
 Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
 Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
 Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,  
 Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugihara, Y.,  
 Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomihaga, N., Toya,  
 T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamanaka, I.,  
 Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,  
 M., Muramatsu, M., and Hayashizaki, Y.  
 TITLE RIKEN Mouse ESTs (Komno, H., et al.)  
 JOURNAL Unpublished  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Saito-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki,  
 N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 Thermostabilization and thermocatalysis of thermolabile enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Kikunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
 Y., and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.



source

1.355  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-BT1-4qp-f-05-0-UI"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="UI-R-BT1"  
/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The library UI-R-BT1 is a subtracted library derived from a mixture of the following tissues: hippocampus, thalamus, mid-brain, medulla, corpus striatum, cerebral cortex and testis. For a detailed description of the library from which this clone was derived, please visit our web site at [ratdesc.eng.uiowa.edu](http://ratdesc.eng.uiowa.edu). The subtraction has been previously described in (Bonaldi, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_SEQ=None found"

BASE COUNT 30 a 126 c 161 g 37 t 1 others

ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 355;  
Best Local Similarity 94.7%; Pred. No. 1.8e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||||  
223 GGGTCTGCAGCGGATGAT 241

Db

RESULT 11  
AA892154 386 bp mRNA linear EST 25-JAN-1999  
DEFINITION EST195957 Normalized rat kidney, Bento Soares Rattus sp. cDNA clone  
AKIAN02.3' end, mRNA sequence.  
AA892154  
VERSION AA892154.1 GI:3019033  
KEYWORDS EST.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 386)  
Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.  
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat Gene Index  
Unpublished  
Contact: Lee, NH  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@igr.org  
Seq primer: M13-21.  
Location/Qualifiers  
1.386  
/organism="Rattus sp."  
/mol\_type="mRNA"  
/db\_xref="ATCC (ihost):2017749"  
/db\_xref="taxon:10116"  
/clone="AKIAN02"  
/clone\_lib="Normalized rat kidney, Bento Soares"  
/note="Organ: kidney; Vector: pRT3D-Pac; Site\_1: EcoRI; Site\_2: NotI"

BASE COUNT 101 a 110 c 96 g 79 t

ORIGIN

Query Match 91.6%; Score 17.4; DB 9; Length 386;  
Best Local Similarity 94.7%; Pred. No. 1.8e+03;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||||  
275 GGGTCTGCAGCGGATGCT 293

Db

RESULT 12  
CB735745/c 463 bp mRNA linear EST 11-APR-2003  
LOCUS AMGNMNC:NRH4-00032-B2-A-W Rat hypothalamus (10464) Rattus  
DEFINITION norvegicus cDNA clone nrh4-00032-b2 5', mRNA sequence.  
CB735745  
VERSION CB735745.1 GI:29802948  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 463)  
Amgen EST Program.  
Amgen Rat EST Program  
Unpublished  
Contact: Dan Fitzpatrick  
Amgen, Inc  
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00032 row: b column: 2.  
Location/Qualifiers  
1.463  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="nrh4-00032-b2"  
/clone\_lib="W Rat hypothalamus (10464)"  
/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; W Rat hypothalamus adult female Wistar rat avg. insert size 2.3 kb fraction 6 and 7"

BASE COUNT 77 a 119 c 103 g 103 t 61 others

ORIGIN

Query Match 91.6%; Score 17.4; DB 14; Length 463;  
Best Local Similarity 94.7%; Pred. No. 1.9e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||||  
259 GGGTCTGCAGCGGATGCT 241

Db

RESULT 13  
CB614996/c 573 bp mRNA linear EST 07-APR-2003  
LOCUS AMGNMNC:NRD1-00012-A1-A nrd1 (10855) Rattus norvegicus cDNA clone  
DEFINITION nrd1-00012-a1 5', mRNA sequence.  
CB614996  
VERSION CB614996.1 GI:29574884  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 573)  
Amgen EST Program.  
Amgen Rat EST Program  
Unpublished  
Contact: Dan Fitzpatrick  
Amgen, Inc  
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00012 row: a column: 1.

FEATURES  
source  
Location/Qualifiers  
1..573  
/organism="Rattus norvegicus"  
/mol\_type="rRNA"  
/db\_xref="taxon:10116"  
/clone="nr01-00012-a1"  
/issue\_type="Dorsal Root Ganglia"  
/clone\_lib="nr01 (10855)"  
/note="Vector: pSPOR1; Site\_1: SalI; Site\_2: NotI; rat dorsal root ganglia"

BASE COUNT 112 a 165 c 149 g 147 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 14; Length 573;  
Best Local Similarity 94.7%; Pred. No. 1.9e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||  
Db 288 GGGTCTGCAGCGGATGCT 270

RESULT 14  
AZ952841/c 580 bp DNA linear GSS 27-APR-2001  
LOCUS 2M021N18 Mouse 10kb plasmid UUGC2M library Mus musculus genomic  
DEFINITION clone UUGC2M021N18 R, genomic survey sequence.  
ACCESSION AZ952841  
VERSION AZ952841.1 GI:13824068  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 580)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weise, R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
Unpublished  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0217 Row: N Column: 18  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 580.  
Location/Qualifiers  
1..580  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M021N18"  
/sex="Female"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC2M library"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114[gb|AF129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 111 a 181 c 156 g 132 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 28; Length 580;  
Best Local Similarity 94.7%; Pred. No. 1.9e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||  
Db 261 GGGTCTGCAGCGGATGCT 243

RESULT 15  
B114889/c 1077 bp mRNA linear EST 26-JUN-2001  
LOCUS 602861478F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:5020578 5',  
DEFINITION mRNA sequence.  
ACCESSION B114889  
VERSION B114889.1 GI:14565790  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1077)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM1837 Row: 1 Column: 19  
High quality sequence stop: 349.  
Location/Qualifiers  
1..1077  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5020578"  
/issue\_type="rhabdomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_17"  
/note="Organ: muscle; Vector: pOTB7; Site\_1: EcoRI; Site\_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 245 a 388 c 340 g 104 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 1077;  
Best Local Similarity 94.7%; Pred. No. 2.1e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy      1 GGCTCTGCAGCGGATGGT 19
        |||||
Db      327 GCGTCTGCAGCTGATGGT 309

RESULT 16
LOCUS   BQ059970/c
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1164)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strusberg, Ph.D.
Email: cga@bs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN).
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LILCM2072 row: h column: 18
High quality sequence start: 19
High quality sequence stop: 418.
Location/Qualifiers
    1..1164
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /clone="IMAGE:5816129"
       /tissue_type="lymphoma, cell line"
       /lab_host="DH10B (phage-resistant)"
       /clone_id="NIH_MGC_99"
       /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
       EcoRI; cDNA made by oligo-dt priming. Directionally cloned
       into EcoRI/XhoI sites using the following 5' adaptor:
       GCCGCGAG(G). Size-selected 500bp for average insert size
       1.8kb. Library constructed by Ling Hong in the laboratory
       of Gerald W. Rubin (University of California, Berkeley)
       using ZAP-cDNA synthesis kit (Stratagene) and Superscript
       II RT (Life Technologies). Note: this is a NIH_MGC
       library."
BASE COUNT      267 a      268 c      307 g      155 t      167 others
ORIGIN
Query Match          91.6%; Score 17.4; DB 12; Length 1164;
Best Local Similarity 94.7%; Pred.No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy      1 GCGTCTGCAGCGGATGGT 19
        |||||
Db      1115 GCGTCTGCAGCGGCTGGT 1097

RESULT 17
LOCUS   BE887786/c
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

Homo sapiens (human)
BE887786
BE887786
BE887786.1 GI:10343430
EST.

```

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 1562)
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LMAG731 row: f column: 15 High quality sequence stop: 6. Location/Qualifiers
FEATURES	1..1562
SOURCE	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:3912590" /tissue_type="leiomyosarcoma" /lab_host="DH10B (phage-resistant)" /clone_1lb="NIH-MGC_71" /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb."
BASE COUNT	567 a 579 c 208 g 208 t
ORIGIN	
Query Match	91.6%; Score 17.4; DB 10; Length 1562;
Best Local Similarity	94.7%; Pred. No. 2.2e+03;
Matches	18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
cy	1 GGGCTCGACGGGATGCT 19 
Db	1152 GGGCTCGACGGGATGCT 1134
RESULT 18	
LOCUS	AKO31524 3874 bp mRNA linear HTC 05-DEC-2002
DEFINITION	Mus musculus 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone:6030447F10 product:hypothetical protein, full insert sequence.
ACCESSION	AKO31524
VERSION	AKO31524.1 GI:26327378
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
JOURNAL	99279253
MEDLINE	10349636
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okaaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
TITLE	
JOURNAL	20493574
MEDLINE	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Sumi, N., Ishi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

**TITLE** RIKEN integrated sequence analysis (RISA) system-384-format genome Res. 10 (11), 1757-1771 (2000)

Qy 2 GGTCTGACGGGATG 18  
 Db 366 GGTCTGACGGGATG 350

## RESULT 20

LOCUS B239259 206 bp mRNA linear EST 17-SEP-2002  
 DEFINITION B239259 Y. Ogihara unpublished cDNA library, wh\_e Triticum  
 accession B239259  
 version B239259.1 GI:23082419  
 keywords EST

SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
 1 Triticaceae; Triticum.  
 1 (bases 1 to 206)  
 Ogihara, Y. and Murai, K.  
 Expressed genes in Triticum aestivum

JOURNAL Contact: Tadao Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp.

## FEATURES

source 1..206  
 Location/Qualifiers

BASE COUNT 35 a 46 c 79 g 44 t 2 others  
 ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 206;  
 Best Local Similarity 94.4%; Pred. No. 4.5e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGTCTGACGGGATG 19  
 Db 89 GGTCTGACGGGATG 106

RESULT 21  
 AM612888 214 bp mRNA linear EST 23-MAR-2000  
 LOCUS AM612888  
 DEFINITION h37g01.x1 NCI\_CGAP\_Col4 Homo sapiens cDNA clone IMAGE:2957328 3',  
 mRNA sequence.

ACCESSION AM612888  
 VERSION AM612888.1 GI:7318074  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 214)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Greg Lemmon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[image.llnl.gov/image/html/liferesources.shtml](http://image.llnl.gov/image/html/liferesources.shtml)  
 Seq primer: -40UP from Gldco.

## FEATURES

source

1..214  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2957328"  
 /tissue\_type="moderately-differentiated adenocarcinoma"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Col4"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.7 kb. Life Technologies catalog #:  
 11531-019"

BASE COUNT 47 a 58 c 71 g 37 t 1 others  
 ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 214;  
 Best Local Similarity 94.4%; Pred. No. 4.5e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGTCTGACGGGATG 18  
 Db 184 GGTCTGACGGGATG 201

## RESULT 22

LOCUS AJ460522/2 240 bp mRNA linear EST 24-MAY-2002  
 DEFINITION AJ460522 S00002 Hordeum vulgare subsp. vulgare cDNA clone  
 S0000200067D02F1, mRNA sequence.

ACCESSION AJ460522  
 VERSION AJ460522.1 GI:21059442  
 KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare  
 ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae; Pooidae  
 1 Triticaceae; Hordeum.  
 1 (bases 1 to 240)  
 Saren, A.-M., Tanskanen, J., Paulin, L. and Schulman, A.H.

JOURNAL Barley EST's  
 Unpublished  
 Contact: Schulman AH  
 Institute of Biotechnology  
 University of Helsinki  
 P.O.Box 56 (Vilkkinkaari 6A), University of Helsinki FIN-00014,  
 Finland.

FEATURES  
 source 1..240  
 Location/Qualifiers

/organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Saana"  
 /db\_xref="taxon:112509"  
 /clone="S0000200067D02F1"  
 /dev\_stage="Embryo"  
 /clone\_lib="S000002"  
 /note="1 day after pollination"

BASE COUNT 47 a 87 c 65 g 41 t  
 ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 240;  
 Best Local Similarity 94.4%; Pred. No. 4.6e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGTCTGACGGGATG 19  
 ||||| ||||| |||||



Db 192 GGCTGAGCGGATGCT 175

RESULT 23  
LOCUS BJ293359/c 253 bp mRNA linear EST 18-SEP-2002  
DEFINITION BJ293359 Y. Ogihara unpublished cDNA library, Wh\_SL Triticum aestivum cDNA clone wh131117 5', mRNA sequence.

ACCESSION BJ293359  
VERSION BJ293359.1 GI:23150883  
KEYWORDS EST  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae  
Triticaceae; Triticum.  
1 (bases 1 to 253)  
Ogihara, Y., and Mural, K.  
Expressed genes in Triticum aestivum  
Unpublished  
Contact: Tadasu Shin-i  
Center for Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tahini@genes.nig.ac.jp.

FEATURES  
SOURCE  
1..253  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="wh131117"  
/issue\_type="seed DPA30"  
/dev\_stage="Feekes" scale 11.3"  
/clone\_1lb="Y. Ogihara unpublished cDNA library, Wh\_SL"  
48 a 90 c 74 g 41 t

BASE COUNT  
ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 253;  
Best Local Similarity 94.4%; Pred. No. 4.6e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGCTGAGCGGATGCT 19  
Db 207 GGCTGAGCGGATGCT 190

RESULT 24  
LOCUS BJ299950 284 bp mRNA linear EST 18-SEP-2002  
DEFINITION BJ299950 Y. Ogihara unpublished cDNA library, Wh\_SL Triticum aestivum cDNA clone wh132119 3', mRNA sequence.

ACCESSION BJ299950  
VERSION BJ299950.1 GI:23154630  
KEYWORDS EST  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae  
Triticaceae; Triticum.  
1 (bases 1 to 284)  
Ogihara, Y., and Mural, K.  
Expressed genes in Triticum aestivum  
Unpublished  
Contact: Tadasu Shin-i  
Center for Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tahini@genes.nig.ac.jp.

FEATURES  
SOURCE  
Location/Qualifiers  
1..284  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="wh132119"  
/issue\_type="seed DPA30"  
/dev\_stage="Feekes" scale 11.3"  
/clone\_1lb="Y. Ogihara unpublished cDNA library, Wh\_SL"  
53 a 67 c 101 g 59 t 4 others

BASE COUNT  
ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 284;  
Best Local Similarity 94.4%; Pred. No. 4.7e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGCTGAGCGGATGCT 19  
Db 148 GGCTGAGCGGATGCT 165

RESULT 25  
LOCUS BJ299963 292 bp mRNA linear EST 18-SEP-2002  
DEFINITION BJ299963 Y. Ogihara unpublished cDNA library, Wh\_SL Triticum aestivum cDNA clone wh132n06 3', mRNA sequence.

ACCESSION BJ299963  
VERSION BJ299963.1 GI:23154638  
KEYWORDS EST  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae  
Triticaceae; Triticum.  
1 (bases 1 to 292)  
Ogihara, Y., and Mural, K.  
Expressed genes in Triticum aestivum  
Unpublished  
Contact: Tadasu Shin-i  
Center for Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tahini@genes.nig.ac.jp.

FEATURES  
SOURCE  
Location/Qualifiers  
1..292  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="wh132n06"  
/issue\_type="seed DPA30"  
/dev\_stage="Feekes" scale 11.3"  
/clone\_1lb="Y. Ogihara unpublished cDNA library, Wh\_SL"  
45 a 82 c 105 g 58 t 2 others

BASE COUNT  
ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 292;  
Best Local Similarity 94.4%; Pred. No. 4.7e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGCTGAGCGGATGCT 19  
Db 34 GGCTGAGCGGATGCT 51

RESULT 26  
LOCUS CA644225 292 bp mRNA linear EST 23-NOV-2002  
DEFINITION CA644225 wrein.pk0078.h1 wrein Triticum aestivum cDNA clone wrein.pk0078.h1 5' end, mRNA sequence.

ACCESSION CA644225  
 VERSION GI:25222521  
 KEYWORDS EST  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
 ; Triticeae; Triticum.  
 1 (bases 1 to 292)  
 Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hailey, C., Yuan, Z.,  
 Miao, G., Caraher, N. and Hanafey, W.K.  
 Miao, G., Caraher, N. and Hanafey, W.K.  
 Dupont wheat cDNA Sequence  
 Unpublished  
 Contact: Scott V. Tingey  
 Crop Genetics  
 E. I. Dupont de Nemours and Company  
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA  
 Tel: 302-631-2602  
 Fax: 302-631-2607  
 Email: Scott.V.Tingey@USA.dupont.com  
 Seq primer: M13.  
 Location/Qualifiers  
 1..292  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4565"  
 /clone="wreln.pk0078.h1"  
 /tissue\_type="root"  
 /clone\_lib="wreln"  
 /note="Vector: pBluescript SK+, Site 1: EcoRI, Site 2:  
 XhoI; Wheat (Triticum aestivum L.) root; normalized from  
 wreln library"  
 BASE COUNT 58 a 67 c 103 g 62 t 2 others  
 ORIGIN  
 Query Match 86.3%; Score 16.4; DB 14; Length 292;  
 Best Local Similarity 94.4%; Pred. No. 4.7e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 2 GGTCTGCAGCGGATGCT 19  
 |||||  
 Db 177 GGTCTGCAGCGGATGCT 194  
 |||||  
 RESULT 27  
 BI013302 324 bp mRNA linear EST 13-JUN-2001  
 LOCUS  
 DEFINITION PM4-ET0154-210101-010-d07 ET0154 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BI013302  
 VERSION BI013302.1 GI:14417373  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 324)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Britones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
 , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922

Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4ct2=PM4-ET0154-  
 210101-010-d07&ct3=2001-01-21&ct4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 13  
 High quality sequence stop: 324.  
 Location/Qualifiers  
 1..324  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="ET0154"  
 /note="Organ: lung tumor; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."  
 BASE COUNT 54 a 83 c 112 g 75 t  
 ORIGIN  
 Query Match 86.3%; Score 16.4; DB 12; Length 324;  
 Best Local Similarity 94.4%; Pred. No. 4.8e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 GGTCTGCAGCGGATGG 18  
 |||||  
 Db 257 GGTCTGCAGCGGATGG 274  
 |||||  
 RESULT 28  
 BU974641 336 bp mRNA linear EST 22-OCT-2002  
 LOCUS  
 DEFINITION HB28J24r BC Hordeum vulgare subsp. vulgare cDNA clone HB28J24  
 5-PRIME, mRNA sequence.  
 ACCESSION BU974641  
 VERSION BU974641.1 GI:24225434  
 KEYWORDS EST.  
 SOURCE Hordeum vulgare subsp. vulgare  
 ORGANISM Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
 ; Triticeae; Hordeum.  
 1 (bases 1 to 336)  
 Radchuk, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.  
 Barley ESTs from developing seeds  
 Unpublished  
 Contact: Stein Nils  
 Molecular Markers Group, Department Genbank  
 Institute of Plant Genetics and Crop Plant Research (IPK)  
 Corrensstr. 3, 06466, Gatersleben, Germany  
 Tel: 039482-5522  
 Fax: 039482-5595  
 Email: stein@ipk-gatersleben.de  
 Insert Length: 336 Std Error: 0.00  
 Plate: 28 row: 0 column: 24  
 Seq primer: M13rev.  
 Location/Qualifiers  
 1..336  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Barke"  
 /db\_xref="GABI:239491"  
 /db\_xref="taxon:112509"  
 /clone="HB28J24"  
 /tissue\_type="developing caryopsis"  
 /dev\_stage="8-15 DAP (days after pollination)"  
 /lab\_host="XL10-Gold"

/clone\_1lb="BC"  
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); developing caryopsis , 8-15 DAP (days after pollination) Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."

BASE COUNT 67 a 122 c 91 g 56 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 13; Length 336;  
Best Local Similarity 94.4%; Pred. No. 4.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCTGCGACGGGATGCT 19  
|||||  
Db 212 GGCTGCGACGGGATGCT 195

RESULT 29 350 bp mRNA linear EST 24-APR-1995  
R25769 Y433f06.f1 Soares Placenta Nb2HP Homo sapiens cDNA clone  
LOCUS R25769.1 GI:781904  
DEFINITION R25769  
VERSION R25769.1 GI:781904  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 350)  
Haller, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hulman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rife, L., Rohlfing, T., Soares, M., Tan, F., Trevaeth, E., Waterston, R., Williamson, A., Woldmann, P. and Wilson, R.  
The Washu-Merck EST Project  
TITLE Unpublished  
JOURNAL Contact: Wilson RK  
COMMENT Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 484  
High quality sequence stops: 220  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lln.gov) for further information.  
Insert Length: 484 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 220.

FEATURES  
source  
1..350  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:538127"  
/db\_xref="taxon:9606"  
/clone="IMAGE:132515"  
/sex="female"  
/dev\_stage="placenta obtained at birth (full term)"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_1lb="Soares Placenta Nb2HP"  
/note="Organ: Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGATTGCGCGCGACGAGATTTTCTTTTCTTTT 3' ], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Library constructed by Bento Soares and M.Fatima Ronaldo. "

BASE COUNT 81 a 107 c 104 g 53 t 5 others

ORIGIN

Query Match 86.3%; Score 16.4; DB 14; Length 350;  
Best Local Similarity 94.4%; Pred. No. 4.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGCGACGGGATGCT 18  
|||||  
Db 89 GGCTGCGACGGGATGCT 72

RESULT 30 355 bp mRNA linear EST 04-MAY-2000  
AM769047  
LOCUS h157h11.x1 NCI CGAP\_C014 Homo sapiens cDNA clone IMAGE:3005349.3  
DEFINITION similar to TR:060997 Q60997 CRP-DUCTIN PRECURSOR ;, mRNA sequence.  
ACCESSION AM769047  
VERSION AM769047.1 GI:7701073  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 355)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL, send email to: info@image.lln.gov  
Possible reversed clone: polyT not found  
Seq primer: -40UP from Gibco  
High quality sequence stop: 225.

FEATURES  
source  
1..355  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3005349"  
/issue\_type="moderately-differentiated adenocarcinoma"  
/lab\_host="DH10B"  
/clone\_1lb="NCI CGAP\_C014"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.7 kb. Life Technologies catalog #:  
11531-019"

BASE COUNT 80 a 101 c 111 g 63 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 355;  
Best Local Similarity 94.4%; Pred. No. 4.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGCGACGGGATGCT 18  
|||||  
Db 184 GGCTGCGACGGGATGCT 201

RESULT 31 362 bp mRNA linear EST 09-MAR-1999  
A1473980  
LOCUS tm04g09.x1 NCI CGAP\_C014 Homo sapiens cDNA clone IMAGE:2155648.3  
DEFINITION



/clone="UI-E-EJ1-ajv-b-09-0-UI"  
/tissue type="fetal eyes, lens, eye anterior segment,  
optic nerve, retina, Retina Foveal and Macular, RPE and  
Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_1lb="UI-E-EJ1"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-EJ1 is a subcloned cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT73-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are: fetal eyes, AGATCAGAGA  
; lens, CGATTAGCGA; eye anterior segment, AATGCCGCGT;  
optic nerve, CCATTAGGTG; retina, CCGCG; Retina Foveal and  
Macular, GTCC; RPE and Choroid, ACCCTA. This library was  
created for the program, Gene Discovery in the Visual  
System, supported by National Eye Institute (NEI)."

## BASE COUNT

97 a 105 c 112 g 66 t

## ORIGIN

Query Match 86.3%; Score 16.4; DB 13; Length 380;  
Best Local Similarity 94.4%; Pred. No. 4.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 GGGTCTGACGCGGATCG 18  
|||||  
304 GGGTCTGACGCGGATCG 287

LOCUS AM662148 382 bp mRNA linear EST\_06-APR-2000  
h120508.x1 NCI CGAP GUI Homo sapiens cDNA clone IMAGE:2972823 3'  
similar to TR:Q60997 Q60997 CRP-DUCTIN PRECURSOR ;, mRNA sequence.  
AM662148  
AM662148.1 GI:7454685  
EST.

ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 382)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: Chris Moskalko, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life  
Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The  
I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
image.llnl.gov/image/html/lifeources.shtml  
Seq primer: -400p from Glbco  
High quality sequence stop: 233.  
Location/Qualifiers

## FEATURES

1..382  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2972823"  
/tissue type="2 pooled high-grade transitional cell  
tumors"

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## SOURCE

1..382  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2972823"  
/tissue type="2 pooled high-grade transitional cell  
tumors"

/lab\_host="DH10B"  
/clone\_1lb="NCI CGAP GUI"  
/note="Organ: genitourinary tract; Vector: PCMV-SPORE;  
Site 1: SalI; Site 2: NotI; Cloned unidirectionally.  
Primer: Oligo dT. Library constructed by Life  
Technologies."

BASE COUNT 88 a 108 c 118 g 68 t

## ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 382;  
Best Local Similarity 94.4%; Pred. No. 4.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 GGGTCTGACGCGGATCG 18  
|||||  
184 GGGTCTGACGCGGATCG 201

LOCUS AA613554 382 bp mRNA linear EST\_31-OCT-1997  
nq22e02.s1 NCI CGAP CO10 Homo sapiens cDNA clone IMAGE:1144634 3'  
mRNA sequence.  
AA613554  
AA613554.1 GI:2464592  
EST.

ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 382)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: Ian Kirsch, M.D., Michael R. Emmert-Buck, M.D.,  
Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 496 Std Error: 0.00  
Seq primer: -40m13 fwd. RT from Amersham.  
Location/Qualifiers

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## SOURCE

1..382  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1144634"  
/tissue type="colon tumor RER+"  
/lab\_host="DH10B"  
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; 1st strand cDNA was prepared from  
RER+ colon tumor, and was then primed with a Not I -  
oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
RI adaptors (Pharmacia), digested with Not I and cloned  
into the Not I and Eco RI sites of the modified pT73  
vector. Library is normalized. Library was constructed by  
Bento Soares and M. Fatima Bonaldo (N-Soares4)."  
BASE COUNT 90 a 98 c 110 g 84 t  
ORIGIN  
Query Match 86.3%; Score 16.4; DB 9; Length 382;  
Best Local Similarity 94.4%; Pred. No. 4.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 GGGTCTGACGCGGATCG 18  
|||||

Db 361 GGGTCGACGCTGATGG 378

RESULT 36

LOCUS BE411862/c 394 bp mRNA linear EST 26-JUL-2000

DEFINITION ISCO10.A09F990908 ITEC ISC Barley Leaf Library Hordeum vulgare subsp. vulgare cDNA clone ISCO10.A09, mRNA sequence.

ACCESSION BE411862

VERSION BE411862.1 GI:9409636

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE 1 (bases 1 to 394)

AUTHORS Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M., Sorrells, M., Watbourn, M., and Wenzel, G.

TITLE International Triticale EST Cooperative (ITEC): Production of Expressed Sequence Tags for Species of the Triticaceae

JOURNAL Unpublished (2000)

COMMENT Contact: Pecchioni N

via San Protasio 302, I-29017 Fiorenzuola d'Arda (PC), ITALY

Tel: 39 0523 983758

Fax: 39 0523 983750

Email: n.pecchioni@iol.it

International Triticale EST Cooperative (ITEC)

http://wheat.pw.usda.gov/genome.

FEATURES

source

1..394

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="Nure"

/db\_xref="taxon:112509"

/clone="ISCO10.A09"

/tissue\_type="leaf"

/dev\_stage="first leaf, cold stress"

/clone\_lib="ITEC ISC Barley Leaf Library"

/note="Vector: UNI-ZAP XR; 0.7-2.0 Kbp average insert size."

BASE COUNT 75 a 143 c 105 g 71 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 394;

Best Local Similarity 94.4%; Pred. No. 4.9e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGTCTGACGCGGATGTT 19

|||||

Db 241 GGTCTGACGCGGATGTT 224

|||||

RESULT 37

LOCUS BM443000/c 400 bp mRNA linear EST 23-JUL-2002

DEFINITION EBR002 SQ002 C13 R root, 3 week, hydroponic grown, low nitrogen, cv Optic\_EBR002 Hordeum vulgare subsp. vulgare cDNA clone

ACCESSION BM443000

VERSION BM443000.2 GI:21947654

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE 1 (bases 1 to 400)

AUTHORS Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticaceae; Hordeum.

FEATURES

1 (bases 1 to 400)

AUTHORS Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L., Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.

TITLE Development of Barley Transcriptome Resources

JOURNAL Unpublished

COMMENT On Feb 1, 2002 this sequence version replaced gi:18473763.

Contact: Waugh R, Marshall DF

Genome Dynamics/Computational Biology

Scottish Crop Research Institute

Invergowrie, Dundee, DD2 5DA, Scotland, UK

Tel: 00 44 1382 562731

Fax: 00 44 1382 562426

Email: est@scri.sari.ac.uk

All sequence has a phred quality score of 20 or over

Seq primer: M13 reverse.

FEATURES

source

1..400

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="Optic"

/db\_xref="taxon:112509"

/clone="EBR002 SQ002\_C13"

/tissue\_type="root"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_lib="root, 3 week, hydroponic grown, low nitrogen, cv Optic, EBR002"

/note="Vector: pSPORT1, Site 1: Sal I; Site 2: Not I; Non-normalised library, directionally cloned into pSPORT1. Derived from roots of 3 week old Nitrogen stressed barley plants. Developed as part of the barley transcriptome resources of BBSRC/SERAD funded cereal IGF (Investigating Gene Function) project."

BASE COUNT 77 a 148 c 104 g 71 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 400;

Best Local Similarity 94.4%; Pred. No. 4.9e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGTCTGACGCGGATGTT 19

|||||

Db 254 GGTCTGACGCGGATGTT 237

|||||

RESULT 38

LOCUS CB871117 402 bp mRNA linear EST 23-APR-2003

DEFINITION HC02F21Y HC Hordeum vulgare subsp. vulgare cDNA clone HC02F21

ACCESSION CB871117

VERSION CB871117.1 GI:30073097

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE 1 (bases 1 to 402)

AUTHORS Varshney, R.K., Zhang, H., Burton, R., Stein, N., Langridge, P. and Graner, A.

Barley ESTs from coleoptile tissue

Unpublished

Contact: Stein Nils

Molecular Markers Group, Department Genbank

Institute of Plant Genetics and Crop Plant Research (IPK)

Corrensstr. 3, 06466, Gatersleben, Germany

Tel: 039482-5522

Fax: 039482-5595

Email: stein@ipk-gatersleben.de

Insert Length: 402 Std Error: 0.00

Plate: 2 row: F column: 21

Seq primer: SP6.

Location/Qualifiers

source

1. .402  
/organism="Hordelum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/strain="cv tokak"  
/db\_xref="GABI:549519"  
/db\_xref="taxon:112509"  
/clone="HC02F21"  
/issue\_type="Root"  
/dev\_stage="3 week old"  
/clone\_1lb="HC"  
/note="6 and 10 hour drought stress by placing plants on moist paper (75% rel. humidity) in light"

BASE COUNT 75 a 106 c 140 g 81 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 14; Length 402;  
Best Local Similarity 94.4%; Pred. No. 4.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGGTCGACGGGATGGT 19  
|||||  
Db 186 GGTCTGAGCGGATGCT 203

RESULT 39  
BH306574/c 403 bp DNA linear GSS 30-NOV-2001  
LOCUS CH230-192M6.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
DEFINITION BH306574  
ACCESSION BH306574  
VERSION BH306574.1 GI:17218982  
KEYWORDS GSS.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 403)  
Zhao, S., Shetty, J., Shatsman, S., Tseng, G., Geer, K., Shvartsbeyn  
'A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, P., de  
Jong, P. and Frazer, C.M.  
Rat BAC End Sequences from Library CHORI-230 EcORI segment  
Unpublished  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rac230.htm). For BAC library  
availability, please contact Pieter de Jong (pdejong@tigr.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/or\_ering/information.htm). BAC end  
page: http://www.tigr.org/cdb/bac\_ends/rac/bac\_end\_intro.html  
Plate: 192 row: M column: 6  
Seq primer: T7  
Class: BAC ends.  
Location/Qualifiers  
1. .403  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SENHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-192M6"  
/sex="Female"  
/cell\_type="Brain"  
/clone\_1lb="CHORI-230 Segment 1"  
/note="Vector: PTARBA2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by  
Pieter de Jong"

BASE COUNT 106 a 108 c 77 g 112 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 28; Length 403;  
Best Local Similarity 94.4%; Pred. No. 4.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGTCGACGGGATGG 18  
|||||  
Db 265 GGGTCGACGAGGATGG 248

RESULT 40  
A1492104 406 bp mRNA linear EST 30-MAR-1999  
LOCUS t907b01.x1 NCI\_CGAP\_CL1 Homo sapiens CDNA IMAGE:2108041.3'  
DEFINITION mRNA sequence.  
A1492104  
ACCESSION A1492104 GI:4393107  
VERSION A1492104.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 406)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,  
M.D., Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
www.bio.lnl.gov/bbrp/image/image.html  
Insert length: 498 Std Error: 0.00  
Seq primer: -40UP from Gibco.  
Location/Qualifiers  
1. .406  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2108041"  
/issue\_type="B-cell, chronic lymphocytic leukemia"  
/lab\_host="DH10B"  
/clone\_1lb="NCI CGAP CL1"  
/note="Vector: pTR73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTACCAATCTAGAGGAGGAGCGGCGATCTTTTCTTTTCTTTTCTTTT  
T 3'] double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pTR73 vector.  
Library is normalized, and was constructed by Bento  
Soares and M. Fatima Bonaldo."

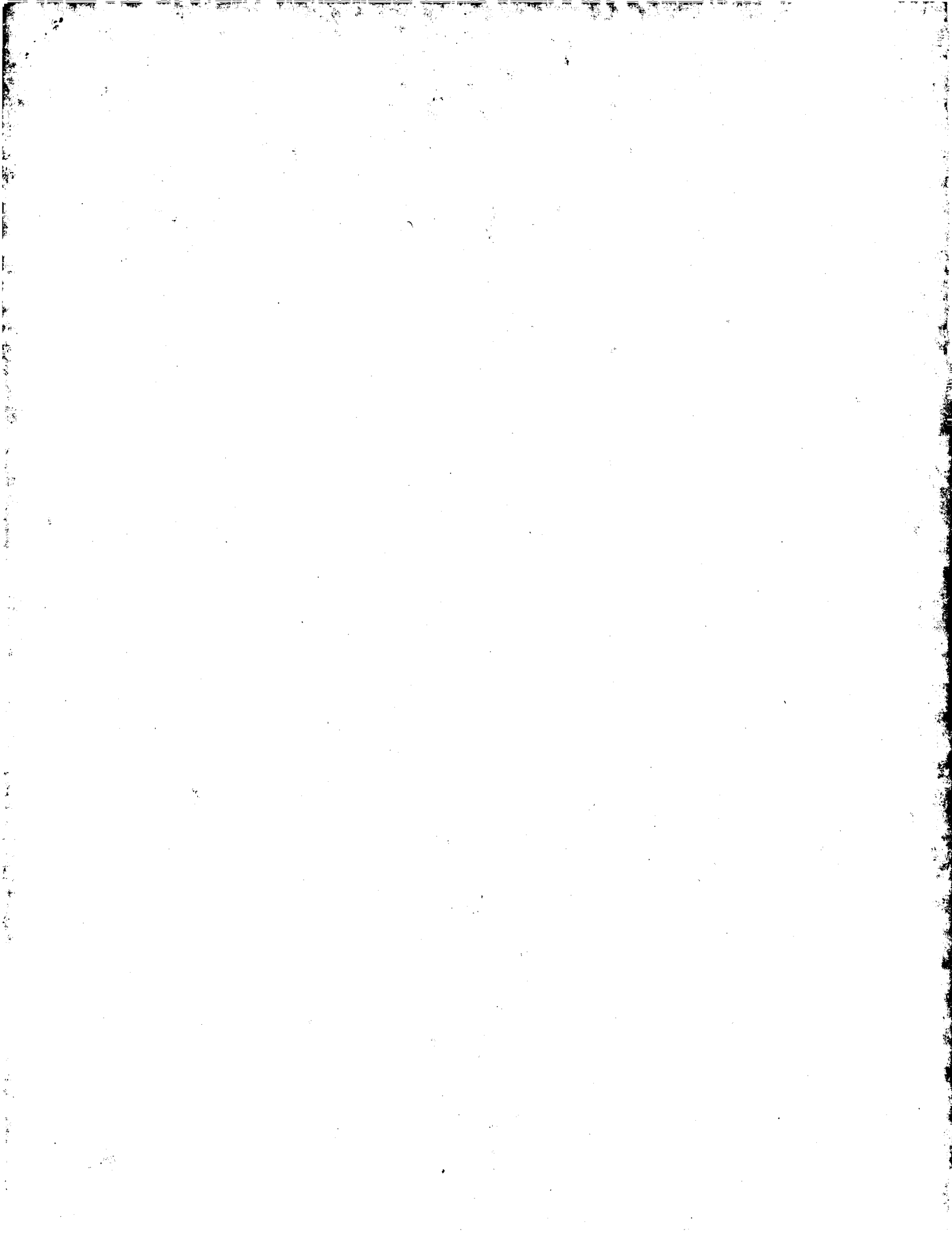
BASE COUNT 94 a 107 c 115 g 90 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 406;  
Best Local Similarity 94.4%; Pred. No. 4.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGTCGACGGGATGG 18  
|||||  
Db 361 GGGTCGACGCTGATGG 378

Search completed: January 13, 2004, 16:33:19  
Job time : 1197.5 secs







```
FILE REFERENCE: 529452000121
CURRENT APPLICATION NUMBER: US/10/085,117
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 361
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 2996
TYPE: DNA
ORGANISM: Homo sapiens
US-10-085-117-11

Query Match      100.0%; Score 19; DB 12; Length 2996;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGCTGCAGCGCGGATGGT 19
        |||
Db      125 GGGCTGCAGCGCGGATGGT 107

RESULT 3
US-10-044-090-546/c
Sequence 546, Application US/10044090
Publication No. US20020137081A1
GENERAL INFORMATION:
APPLICANT: Olga Bandman
TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
FILE REFERENCE: PA-0028 US
CURRENT APPLICATION NUMBER: US/10/044,090
CURRENT FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 850
SOFTWARE: PERL Program
SEQ ID NO 546
LENGTH: 3439
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20020137081A1 3420843CB1
US-10-044-090-546

Query Match      100.0%; Score 19; DB 14; Length 3439;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGCTGCAGCGCGGATGGT 19
        |||
Db      273 GGGCTGCAGCGCGGATGGT 255

RESULT 4
US-10-085-117-10/c
Sequence 10, Application US/10085117
Publication No. US2003023234A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: 529452000121
CURRENT APPLICATION NUMBER: US/10/085,117
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 361
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 36595
TYPE: DNA
ORGANISM: Homo sapiens
US-10-085-117-10
```

```
Query Match      100.0%; Score 19; DB 12; Length 36595;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGCTGCAGCGCGGATGGT 19
        |||
Db      11085 GGGCTGCAGCGCGGATGGT 11067

RESULT 5
US-09-918-995-7315/c
Sequence 7315, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hysed, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7315
LENGTH: 390
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(390)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-7315

Query Match      86.3%; Score 16.4; DB 11; Length 390;
Best Local Similarity 94.4%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GGGCTGCAGCGCGGATGG 18
        |||
Db      118 GGGCTGCAGCTGATGG 101

RESULT 6
US-10-183-708-49/c
Sequence 49, Application US/10183708
Publication No. US20030143679A1
GENERAL INFORMATION:
APPLICANT: VOSSHALL, LESLIE
APPLICANT: AMREIN, HUBERT
APPLICANT: AXEL, RICHARD
TITLE OF INVENTION: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF
FILE REFERENCE: 0575/58715-AA-PCT-US/JPM/ADM/BJA
CURRENT APPLICATION NUMBER: US/10/183,708
CURRENT FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 09/932,227
PRIOR FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: PCT/US00/04995
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/257,706
PRIOR FILING DATE: 1999-02-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: PatentIn version 3.0
SEQ ID NO 49
LENGTH: 1194
TYPE: DNA
ORGANISM: Drosophila melanogaster DOR119
US-10-183-708-49

Query Match      86.3%; Score 16.4; DB 13; Length 1194;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 2 GGCTGCAGCGGATGCT 19  
Db 933 GGCTGCAGCGGATGCT 916

## RESULT 7

US-09-932-227-49/c  
; Sequence 49, Application US/09932227  
; Publication No. US20030186359A1  
; GENERAL INFORMATION:  
; APPLICANT: VOSSHALL, LESLIE  
; APPLICANT: AMREIN, HUBERT  
; APPLICANT: AXEL, RICHARD  
; TITLE OF INVENTION: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF  
; FILE REFERENCE: 0575/58715-A-PCT-US/JPW/ADM/BJA  
; CURRENT APPLICATION NUMBER: US/09/932,227  
; PRIOR FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: PCT/US00/04995  
; PRIOR FILING DATE: 2000-02-25  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 49  
; LENGTH: 1194  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster DOR119  
US-09-932-227-49

Query Match 86.3%; Score 16.4; DB 13; Length 1194;  
Best Local Similarity 94.4%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGCTGCAGCGGATGCT 19  
Db 933 GGCTGCAGCGGATGCT 916

## RESULT 8

US-09-738-626-109  
; Sequence 109, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 109  
; LENGTH: 1824  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-109

Query Match 86.3%; Score 16.4; DB 10; Length 1824;  
Best Local Similarity 94.4%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGCTGCAGCGGATGCT 19  
Db 301 GGCTGCAGCGGATGCT 318

## RESULT 9

US-09-738-626-1/c  
; Sequence 1, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 1  
; LENGTH: 3309400  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-1

Query Match 86.3%; Score 16.4; DB 10; Length 3309400;  
Best Local Similarity 94.4%; Pred. No. 43;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGCTGCAGCGGATGCT 19  
Db 103135 GGCTGCAGCGGATGCT 103118

## RESULT 10

US-10-029-386-14644  
; Sequence 14644, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 14644  
; LENGTH: 153  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR22 166,0  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.6  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.71  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72  
OTHER INFORMATION: SWISSPROT HIT: P33927, EVALUATE 2.00e-16

OTHER INFORMATION: NT HIT: G11418060, EVALUE 6.00e-50  
OTHER INFORMATION: EST\_HUMAN HIT: BF884753.1, EVALUE 1.00e-27  
US-10-029-386-14644

Query Match 83.2%; Score 15.8; DB 13; Length 153;  
Best Local Similarity 89.5%; Pred. No. 3.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 19  
|||||  
Db 133 GGGTCTGCAGCAGCGGT 151

## RESULT 11

US-10-027-632-6525/C  
Sequence 6525, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6525  
LENGTH: 504  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-6525

Query Match 83.2%; Score 15.8; DB 13; Length 504;  
Best Local Similarity 89.5%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 19  
|||||  
Db 493 GGGTCTGCAGCAGGTTGT 475

## RESULT 12

US-10-027-632-6525/C  
Sequence 6525, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6525  
LENGTH: 504  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-6525

Query Match 83.2%; Score 15.8; DB 14; Length 504;  
Best Local Similarity 89.5%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 19  
|||||  
Db 493 GGGTCTGCAGCAGGTTGT 475

RESULT 13  
US-10-027-632-188580  
Sequence 188580, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 188580  
LENGTH: 507  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-188580

Query Match 83.2%; Score 15.8; DB 13; Length 507;  
Best Local Similarity 89.5%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 19  
|||||  
Db 18 GGGTCTGCAGCAGGTTGT 36

RESULT 14  
US-10-027-632-188581  
Sequence 188581, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24

```
FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 188581
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-188581
```

```
Query Match      83.2% Score 15.8; DB 13; Length 507;
Best Local Similarity 89.5%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GGGTCGACGCGGATGCT 19
        |||||
Db      18 GGGTCGACGCGGATGCT 36
```

```
RESULT 15
US-10-027-632-188582
; Sequence 188582, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 188582
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-188582
```

```
Query Match      83.2% Score 15.8; DB 13; Length 507;
Best Local Similarity 89.5%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GGGTCGACGCGGATGCT 19
        |||||
```

```
Db      18 GGGTCGACGCGGATGCT 36
```

```
RESULT 16
US-10-027-632-188583
; Sequence 188583, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 188583
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-188583
```

```
Query Match      83.2% Score 15.8; DB 13; Length 507;
Best Local Similarity 89.5%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GGGTCGACGCGGATGCT 19
        |||||
Db      18 GGGTCGACGCGGATGCT 36
```

```
RESULT 17
US-10-027-632-188580
; Sequence 188580, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 188580
; LENGTH: 507
```

TYPE: DNA  
ORGANISM: Human  
US-10-027-632-188580

Query Match 83.2%; Score 15.8; DB 14; Length 507;  
Best Local Similarity 89.5%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCTGCAGCGGATGCT 19  
|||||  
Db 18 GGGCTGCAGCGGATGCT 36

RESULT 18  
US-10-027-632-188581  
Sequence 188581, Application US/10027632

GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME

FILE REFERENCE: 108827.129  
CURRENT FILING DATE: 2002-04-30  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 188581  
LENGTH: 507  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-188581

Query Match 83.2%; Score 15.8; DB 14; Length 507;  
Best Local Similarity 89.5%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCTGCAGCGGATGCT 19  
|||||  
Db 18 GGGCTGCAGCGGATGCT 36

RESULT 19  
US-10-027-632-188582  
Sequence 188582, Application US/10027632

GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME

FILE REFERENCE: 108827.129  
CURRENT FILING DATE: 2002-04-30  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 188582  
LENGTH: 507  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-188582

QY 1 GGGCTGCAGCGGATGCT 19  
|||||  
Db 18 GGGCTGCAGCGGATGCT 36

Query Match 83.2%; Score 15.8; DB 14; Length 507;  
Best Local Similarity 89.5%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCTGCAGCGGATGCT 19  
|||||  
Db 18 GGGCTGCAGCGGATGCT 36

RESULT 20  
US-10-027-632-188583  
Sequence 188583, Application US/10027632

GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME

FILE REFERENCE: 108827.129  
CURRENT FILING DATE: 2002-04-30  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 188583  
LENGTH: 507  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-188583

Query Match 83.2%; Score 15.8; DB 14; Length 507;  
Best Local Similarity 89.5%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCTGCAGCGGATGCT 19  
|||||  
Db 18 GGGCTGCAGCGGATGCT 36

RESULT 21  
US-10-029-386-941  
Sequence 941, Application US/10029386

PUBLICATION NO. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE  
EXPRESSION ANALYSIS TWO

```
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 941
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR22-166.0
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.71
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72
; OTHER INFORMATION: NT HIT: g11418060, EVALUE 2.00e-50
; OTHER INFORMATION: EST HUMAN HIT: BF884753.1, EVALUE 3.00e-28
; OTHER INFORMATION: SWISSPROT HIT: P32927, EVALUE 3.00e-16
US-10-029-386-941
```

```
Query Match      83.2%; Score 15.8; DB 13; Length 589;
Best Local Similarity 89.5%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GGGTCTGCAGCGGATGCT 19
         |||||
Db      412 GGGTCTGCAGCGGATGCT 430
```

```
RESULT 22
US-10-027-632-159729/c
; Sequence 159729, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159729
; LENGTH: 868
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-159729
```

```
Query Match      83.2%; Score 15.8; DB 13; Length 868;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GGGTCTGCAGCGGATGCT 19
         |||||
Db      372 GGGTCTGCAGCGGATGCT 354
```

```
RESULT 23
US-10-027-632-159729/c
```

```
; Sequence 159729, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159729
; LENGTH: 868
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-159729
```

```
Query Match      83.2%; Score 15.8; DB 14; Length 868;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GGGTCTGCAGCGGATGCT 19
         |||||
Db      372 GGGTCTGCAGCGGATGCT 354
```

```
RESULT 24
US-10-027-632-174763/c
; Sequence 174763, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174763
; LENGTH: 2940917
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(2940917)
```

OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-174763

Query Match 83.2%; Score 15.8; DB 13; Length 2940917;  
Best Local Similarity 89.5%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
DB 299508 GGGTATTCAGCGGATGCT 299490

RESULT 25  
US-10-027-632-174763/C  
Sequence 174763, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 174763  
LENGTH: 2940917  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(2940917)  
OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-174763

Query Match 83.2%; Score 15.8; DB 14; Length 2940917;  
Best Local Similarity 89.5%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
DB 299508 GGGTATTCAGCGGATGCT 299490

RESULT 26  
US-10-369-493-44678  
Sequence 44678, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2003-02-28  
NUMBER OF SEQ ID NOS: 60/360,039  
SOFTWARE: Friedmann macro Napro4

PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 44678  
LENGTH: 1560  
TYPE: DNA  
ORGANISM: SYNECHOCOCCUS SP. WH 8102  
US-10-369-493-44678

Query Match 81.1%; Score 15.4; DB 12; Length 1560;  
Best Local Similarity 94.1%; Pred. No. 3.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GTCTGCAGCGGATGCT 19  
DB 1373 GTCTGCAGCGGATGCT 1389

RESULT 27  
US-09-939-964-1/C  
Sequence 1, Application US/09939964  
Publication No. US20030054522A1  
GENERAL INFORMATION:  
APPLICANT: Rosenthal, Andre  
APPLICANT: Freiberg, Christoph  
APPLICANT: Perret, Xavier Philippe  
APPLICANT: Broughton, William John  
TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic  
FILE REFERENCE: CAP0068  
CURRENT APPLICATION NUMBER: US/09/939,964  
CURRENT FILING DATE: 2001-08-27  
PRIOR APPLICATION NUMBER: 09/214,808  
PRIOR FILING DATE: 1999-06-22  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 536165  
TYPE: DNA  
ORGANISM: Rhizobium  
US-09-939-964-1

Query Match 81.1%; Score 15.4; DB 11; Length 536165;  
Best Local Similarity 94.1%; Pred. No. 1.7e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GTCTGCAGCGGATGCT 19  
DB 457877 GTCTGCAGCGGATGCT 457861

RESULT 28  
US-09-818-875-1345  
Sequence 1345, Application US/09818875  
Publication No. US20030051270A1  
GENERAL INFORMATION:  
APPLICANT: Kmiec, Eric B.  
APPLICANT: Gamper, Howard B.  
APPLICANT: Rice, Michael C.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/09/818,875  
CURRENT FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
PRIOR FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 4385  
SOFTWARE: Friedmann macro Napro4



SEQ ID NO 1345  
LENGTH: 121  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-818-875-1345

Query Match 78.9%; Score 15; DB 11; Length 121;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGA 15  
DB 89 GGGTCTGCAGCGGA 103

RESULT 29  
US-09-818-875-1346/C  
Sequence 1346, Application US/09818875  
Publication No. US20030051270A1  
GENERAL INFORMATION:  
APPLICANT: Kmiec, Eric B.  
APPLICANT: Gamper, Howard B.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/09/818,875  
CURRENT FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
PRIOR FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 4385  
SOFTWARE: Friedman macro Napro4  
SEQ ID NO 1346  
LENGTH: 121  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-818-875-1346

Query Match 78.9%; Score 15; DB 11; Length 121;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGA 15  
DB 33 GGGTCTGCAGCGGA 19

RESULT 30  
US-10-209-787-1345  
Sequence 1345, Application US/10209787  
Publication No. US20030217377A1  
GENERAL INFORMATION:  
APPLICANT: Kmiec, Eric B.  
APPLICANT: Gamper, Howard B.  
APPLICANT: Rice, Michael C.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/10/209,787  
CURRENT FILING DATE: 2002-07-30  
PRIOR APPLICATION NUMBER: US 09/818,875  
PRIOR FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538

PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
PRIOR FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 4385  
SOFTWARE: Friedman macro Napro4  
SEQ ID NO 1345  
LENGTH: 121  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-209-787-1345

Query Match 78.9%; Score 15; DB 13; Length 121;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGA 15  
DB 89 GGGTCTGCAGCGGA 103

RESULT 31  
US-10-209-787-1346/C  
Sequence 1346, Application US/10209787  
Publication No. US20030217377A1  
GENERAL INFORMATION:  
APPLICANT: Kmiec, Eric B.  
APPLICANT: Gamper, Howard B.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/10/209,787  
CURRENT FILING DATE: 2002-07-30  
PRIOR APPLICATION NUMBER: US 09/818,875  
PRIOR FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
PRIOR FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 4385  
SOFTWARE: Friedman macro Napro4  
SEQ ID NO 1346  
LENGTH: 121  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-209-787-1346

Query Match 78.9%; Score 15; DB 13; Length 121;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGA 15  
DB 33 GGGTCTGCAGCGGA 19

RESULT 32  
US-10-429-802-28  
Sequence 28, Application US/10429802  
Publication No. US20030228285A1  
GENERAL INFORMATION:  
APPLICANT: HUNG, MIEN-CHIE  
APPLICANT: WONG, KA YIN  
APPLICANT: ZOU, YIYU  
TITLE OF INVENTION: BIPARTITE T-CELL FACTOR (TCF)-RESPONSIVE PROMOTER  
FILE REFERENCE: UTSC:7520S  
CURRENT APPLICATION NUMBER: US/10/429,802  
CURRENT FILING DATE: 2003-05-05  
PRIOR APPLICATION NUMBER: 60/377,672

PRIOR FILING DATE: 2002-05-03  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 28  
LENGTH: 340  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-429-802-28

Query Match 78.9%; Score 15; DB 12; Length 340;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GGGTCTGCAGCGGGA 15  
|||||  
Db 264 GGGTCTGCAGCGGGA 278

## RESULT 33

US-09-016-869B-22  
Sequence 22, Application US/09016869B  
Patent No. US20020082392A1  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Demetrick, Douglas J.  
APPLICANT: Serrano, Manuel  
APPLICANT: Hannon, Gregory J.  
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and  
TITLE OF INVENTION: Uses Related Thereto  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ropes & Gray  
STREET: One International Place  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordpad  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,869B  
FILING DATE: 30-JAN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/893,274  
FILING DATE: 15-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,511  
FILING DATE: 14-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,812  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,371  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,915  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,997  
FILING DATE: 17-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: GPCI-P10-071  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 951-7739  
TELEFAX: (617) 951-7050  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 401 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-016-869B-22

Query Match 78.9%; Score 15; DB 9; Length 401;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GGGTCTGCAGCGGGA 15  
|||||  
Db 18 GGGTCTGCAGCGGGA 32

## RESULT 34

US-09-016-869B-21  
Sequence 21, Application US/09016869B  
Patent No. US20020082392A1  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Demetrick, Douglas J.  
APPLICANT: Serrano, Manuel  
APPLICANT: Hannon, Gregory J.  
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and  
TITLE OF INVENTION: Uses Related Thereto  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ropes & Gray  
STREET: One International Place  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordpad  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,869B  
FILING DATE: 30-JAN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/893,274  
FILING DATE: 15-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,511  
FILING DATE: 14-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,812  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,371  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,915  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,997  
FILING DATE: 17-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: GPCI-P10-071  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 951-7739  
TELEFAX: (617) 951-7050  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 404 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: CDNA  
US-09-016-869B-21

Query Match 78.9%; Score 15; DB 9; Length 404;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGCTGCGAGCGGA 15  
|||||  
Db 112 GGGCTGCGAGCGGA 126

## RESULT 35

US-09-908-671-1  
Sequence 1, Application US/09908671  
Publication No. US20030138928A1

GENERAL INFORMATION:

APPLICANT: CARSON, DENNIS A.

NOBORI, TSUTOMU

TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR  
DETECTION OF CANCER, MONITORING OF TUMOR PROGRESSION AND CA  
TREATMENT

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/908,671

FILING DATE: 18-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/227,800

FILING DATE: 14-Apr-1994

ATTORNEY/AGENT INFORMATION:

NAME: HOWELLS, STACY L.

REGISTRATION NUMBER: 34,842

REFERENCE/DOCKET NUMBER: 07340/023001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1146 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

IMMEDIATE SOURCE:

LIBRARY: 5'-region of CDK4I (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1146

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-908-671-1

Query Match 78.9%; Score 15; DB 13; Length 1146;

Best Local Similarity 100.0%; Pred. No. 5.8e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGCTGCGAGCGGA 15  
|||||  
Db 525 GGGCTGCGAGCGGA 539

## RESULT 36

US-09-984-664-15

Sequence 15, Application US/09984664

Publication No. US2003009950A1

GENERAL INFORMATION:

APPLICANT: Hanna, Michelle

TITLE OF INVENTION: Molecular Detection Systems Utilizing Retrievable Oligonucleotide

FILE REFERENCE: 2072, 0010000

CURRENT APPLICATION NUMBER: US/09/984,664

CURRENT FILING DATE: 2001-10-30

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentin version 3.0

SEQ ID NO 15

LENGTH: 3621

TYPE: DNA

ORGANISM: Homo sapiens

US-09-984-664-15

Query Match 78.9%; Score 15; DB 11; Length 3621;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGCTGCGAGCGGA 15  
|||||  
Db 2378 GGGCTGCGAGCGGA 2392

## RESULT 37

US-09-993-525-11

Sequence 11, Application US/09993525

Publication No. US20030032093A1

GENERAL INFORMATION:

APPLICANT: Diaz-Torres, Maria

TITLE OF INVENTION: Production of Secreted Polypeptides

FILE REFERENCE: GC618

CURRENT APPLICATION NUMBER: US/09/993,525

CURRENT FILING DATE: 2001-11-14

PRIOR APPLICATION NUMBER: WO PCT/US99/31010

PRIOR FILING DATE: 1999-12-21

NUMBER OF SEQ ID NOS: 17

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11

LENGTH: 52

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: primer

US-09-993-525-11

Query Match 77.9%; Score 14.8; DB 11; Length 52;

Best Local Similarity 88.9%; Pred. No. 1.1e+03;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGCTGCGAGCGGATG 18  
|||||  
Db 3 GCGCTGCGAGCGGATG 20

## RESULT 38

US-10-156-761-7227/c

Sequence 7227, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

```

; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7227
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(219)
US-10-156-761-7227
```

```

Query Match      77.9%; Score 14.8; DB 15; Length 219;
Best Local Similarity 88.9%; Pred. No. 9,1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GGGTCTGCGCGGCGATGG 18
         |||||
Db      103 GGGTCGCGCGGCGCTGG 86
```

```

RESULT 39
US-10-339-604-69/c
; Sequence 69, Application US/10339604
; Publication No. US20030152982A1
; GENERAL INFORMATION:
; APPLICANT: DE BEENHOWER, HANS
; APPLICANT: PORTAELS, FRANCOISE
; APPLICANT: MACHTELINCKX, LIEVE
; APPLICANT: JANNES, GERT
; APPLICANT: ROSSAU, RUDI
; TITLE OF INVENTION: Oligonucleotide Molecules for Use in Detection of Mycobacterium
; FILE REFERENCE: 1657.0010001
; CURRENT APPLICATION NUMBER: US/10/339,604
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: US/09/722,319
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 08/750,088
; PRIOR FILING DATE: 1996-12-06
; PRIOR APPLICATION NUMBER: PCT/EP95/02230
; PRIOR FILING DATE: 1995-06-09
; PRIOR APPLICATION NUMBER: EP 94870093.5
; PRIOR FILING DATE: 1994-06-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Probe
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: n can be any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (33)..(33)
; OTHER INFORMATION: n can be any nucleotide
US-10-339-604-69
```

```

Query Match      77.9%; Score 14.8; DB 13; Length 228;
Best Local Similarity 88.9%; Pred. No. 9e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GGGTCTGCGCGGCGATGG 18
         |||||
Db      76 GGGTCTGCGCGGCGATGG 59
```

```

RESULT 40
US-09-878-574-12285/c
; Sequence 12285, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(115401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 12285
; LENGTH: 276
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701065580H1
US-09-878-574-12285
```

```

Query Match      77.9%; Score 14.8; DB 10; Length 276;
Best Local Similarity 88.9%; Pred. No. 8.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GGGTCTGCGCGGCGATGG 18
         |||||
Db      59 GGGTCTGCATCGGGAAGG 42
```

Search completed: January 13, 2004, 21:03:11  
Job time : 139.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 13, 2004, 15:53:18 ; Search time 36.5 Seconds  
(without alignments)  
229.761 Million cell updates/sec

Title: US-09-719-737-9  
Perfect score: 19  
Sequence: 1 99gctcgcagcg9atgct 19

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19	100.0	3475	1 US-07-960-389-1	Sequence 1, Appli
C 2	16.4	86.3	1350	4 US-09-252-991A-4370	Sequence 4370, Ap
C 3	16.4	86.3	1455	4 US-09-252-991A-4216	Sequence 4216, Ap
C 4	16.4	86.3	5802	4 US-09-341-587-4	Sequence 4, Appli
C 5	16.4	86.3	28720	4 US-09-341-587-7	Sequence 7, Appli
C 6	15.4	81.1	480	4 US-09-252-991A-3819	Sequence 3819, Ap
C 7	15.4	81.1	897	4 US-09-252-991A-3775	Sequence 3775, Ap
C 8	15.4	81.1	1272	4 US-09-252-991A-3591	Sequence 3591, Ap
C 9	15.4	81.1	536165	4 US-09-214-808-1	Sequence 1, Appli
C 10	15.4	81.1	4403765	3 US-09-103-840A-2	Sequence 2, Appli
C 11	15.4	81.1	4411529	3 US-09-103-840A-1	Sequence 1, Appli
C 12	15	78.9	401	3 US-08-581-918A-27	Sequence 27, Appli
C 13	15	78.9	401	3 US-08-346-147B-27	Sequence 27, Appli
C 14	15	78.9	401	4 US-08-497-214D-27	Sequence 27, Appli
C 15	15	78.9	404	3 US-08-581-918A-26	Sequence 26, Appli
C 16	15	78.9	404	3 US-08-346-147B-26	Sequence 26, Appli
C 17	15	78.9	404	4 US-08-497-214D-26	Sequence 26, Appli
C 18	15	78.9	1149	1 US-08-474-177-3	Sequence 26, Appli
C 19	15	78.9	1149	1 US-08-487-033-3	Sequence 3, Appli
C 20	15	78.9	1149	1 US-08-480-810-3	Sequence 3, Appli
C 21	15	78.9	1149	2 US-08-508-735-3	Sequence 3, Appli
C 22	15	78.9	1149	2 US-08-848-251-3	Sequence 3, Appli
C 23	15	78.9	1149	2 US-08-486-047-3	Sequence 3, Appli
C 24	15	78.9	1149	3 US-09-120-130-3	Sequence 3, Appli
C 25	15	78.9	1149	3 US-09-115-252-3	Sequence 3, Appli
C 26	15	78.9	1149	3 US-08-986-515-3	Sequence 3, Appli
C 27	15	78.9	1149	3 US-09-120-128-3	Sequence 3, Appli

C 28	15	78.9	1149	3 US-09-120-129-3	Sequence 3, Appli
C 29	15	78.9	1149	3 US-09-201-139-3	Sequence 3, Appli
C 30	15	78.9	1149	3 US-09-120-131-3	Sequence 3, Appli
C 31	15	78.9	36941	4 US-08-311-731A-130	Sequence 110, App
C 32	14.8	77.9	52	4 US-09-595-386-11	Sequence 11, Appli
C 33	14.8	77.9	228	4 US-08-750-088A-69	Sequence 69, Appli
C 34	14.8	77.9	306	3 US-09-147-935A-5	Sequence 5, Appli
C 35	14.8	77.9	306	3 US-09-147-935A-11	Sequence 11, Appli
C 36	14.8	77.9	306	3 US-09-147-935A-17	Sequence 17, Appli
C 37	14.8	77.9	306	3 US-09-147-935A-18	Sequence 18, Appli
C 38	14.8	77.9	306	3 US-09-147-935A-22	Sequence 22, Appli
C 39	14.8	77.9	306	3 US-09-147-935A-28	Sequence 28, Appli
C 40	14.8	77.9	306	3 US-09-147-935A-29	Sequence 29, Appli
C 41	14.8	77.9	306	3 US-09-147-935A-35	Sequence 35, Appli
C 42	14.8	77.9	306	3 US-09-147-935A-38	Sequence 38, Appli
C 43	14.8	77.9	319	4 US-08-750-088A-35	Sequence 35, Appli
C 44	14.8	77.9	324	4 US-08-750-088A-36	Sequence 36, Appli
C 45	14.8	77.9	418	4 US-09-072-596-302	Sequence 302, App

## ALIGNMENTS

RESULT 1  
US-07-960-389-1/c  
Sequence 1, Application US/07960389  
Patent No. 5705611  
GENERAL INFORMATION:  
APPLICANT: HAYASHIDA, Kaunhiro;  
TITLE OF INVENTION: Human GM-CSF Receptor Component  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schering-Plough Corporation  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disc  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: System Software 7.1  
SOFTWARE: Microsoft Word 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/960,389  
FILING DATE: 07-JAN-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 554,745  
FILING DATE: 18-JUL-1990  
APPLICATION NUMBER: PCT/US 91/04846  
FILING DATE: 16-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Blasdale, John H. C.  
REGISTRATION NUMBER: 31,895  
REFERENCE/DOCKET NUMBER: DX01430  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2902  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3475 base pairs  
TYPE: nucleotides  
STRANDEDNESS: single  
TOPOLOGY: linear  
ORIGINAL SOURCE:  
ORGANISM: human  
FEATURE:  
OTHER INFORMATION: DNA sequence encoding Human GM-CSF receptor  
US-07-960-389-1  
Query Match 100.0%; Score 19; DB 1; Length 3475;  
Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 19

Db 297 GGGTCTGCAGCGGATGCT 279

## RESULT 2

US-09-252-991A-4370/C

; Sequence 4370, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4370

; LENGTH: 1350

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4370

Query Match

Best Local Similarity 86.3%; Score 16.4; DB 4; Length 1350;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGTCTGCAGCGGATGCT 19

Db 783 GGGCTGCAGCGGATGCT 766

## RESULT 3

US-09-252-991A-4216

; Sequence 4216, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4216

; LENGTH: 1455

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4216

## Query Match

Best Local Similarity 86.3%; Score 16.4; DB 4; Length 1455;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGTCTGCAGCGGATGCT 19

Db 709 GGGCTGCAGCGGATGCT 726

## RESULT 4

US-09-341-587-4/C

; Sequence 4, Application US/09341587

; Patent No. 6346606

; GENERAL INFORMATION:

; APPLICANT: Mollenhauer, Jan

; TITLE OF INVENTION: Protein Containing an SRGR Domain

; FILE REFERENCE: 4121-108

; CURRENT FILING DATE: 1999-08-31

; PRIOR FILING DATE: 1998-01-09

; PRIOR APPLICATION NUMBER: PCT/DE98/00096

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 4

; LENGTH: 5802

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-341-587-4

Query Match

Best Local Similarity 86.3%; Score 16.4; DB 4; Length 5802;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 18

Db 5435 GGGTCTGCAGCGGATGCT 5418

## RESULT 5

US-09-341-587-7/C

; Sequence 7, Application US/09341587

; Patent No. 6346606

; GENERAL INFORMATION:

; APPLICANT: Mollenhauer, Jan

; TITLE OF INVENTION: Protein Containing an SRGR Domain

; FILE REFERENCE: 4121-108

; CURRENT FILING DATE: 1999-08-31

; PRIOR FILING DATE: 1998-01-09

; PRIOR APPLICATION NUMBER: PCT/DE98/00096

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 7

; LENGTH: 28720

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-341-587-7

Query Match

Best Local Similarity 86.3%; Score 16.4; DB 4; Length 28720;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 18

Db 28273 GGGTCTGCAGCGGATGCT 28256

## RESULT 6

US-09-252-991A-3819/C

; Sequence 3819, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 3819

; LENGTH: 480

; TYPE: DNA

ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3819

Query Match  
Best Local Similarity 81.1%; Score 15.4; DB 4; Length 480;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCTGCAGCGGATGG 18  
|||||  
Db 212 GGCTGCAGCGGATGG 196

RESULT 7  
US-09-252-991A-3775/c  
; Sequence 3775, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 3775  
; LENGTH: 897  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3775

Query Match  
Best Local Similarity 81.1%; Score 15.4; DB 4; Length 897;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCTGCAGCGGATGG 18  
|||||  
Db 681 GGCTGCAGCGGATGG 665

RESULT 8  
US-09-252-991A-3591  
; Sequence 3591, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 3591  
; LENGTH: 1272  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3591

Query Match  
Best Local Similarity 81.1%; Score 15.4; DB 4; Length 1272;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCTGCAGCGGATGG 18  
|||||  
Db 679 GGCTGCAGCGGATGG 695

RESULT 9  
US-09-214-808-1/c  
; Sequence 1, Application US/09214808A  
; Patent No. 6475793  
; GENERAL INFORMATION:  
; APPLICANT: Rosenthal, Andre  
; APPLICANT: Freiberg, Christoph  
; APPLICANT: Perret, Xavier Philippe  
; APPLICANT: Broughton, William John  
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic  
; Patent No. 6475793  
; TITLE OF INVENTION: Plasmid  
; FILE REFERENCE: CARP0068  
; CURRENT APPLICATION NUMBER: US/09/214,808A  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: PCT/IB97/00950  
; PRIOR FILING DATE: 1997-07-10  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 536165  
; TYPE: DNA  
; ORGANISM: Rhizobium  
US-09-214-808-1

Query Match  
Best Local Similarity 81.1%; Score 15.4; DB 4; Length 536165;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GTCTGCAGCGGATGCT 19  
|||||  
Db 457877 GTCTGCAGCGGATGCT 457861

RESULT 10  
US-09-103-840A-2  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; PRIOR FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match  
Best Local Similarity 81.1%; Score 15.4; DB 3; Length 4403765;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGCAGCGGATG 17  
|||||  
Db 4102066 GGCTGCAGCGGATG 4102082

RESULT 11  
US-09-103-840A-1  
; Sequence 1, Application US/09103840A

Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 81.1%; Score 15.4; DB 3; Length 4411529;  
Best Local Similarity 94.1%; Pred. No. 85;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGGTCTGCAGCGGATG 17  
DB 4109869 GGGTCTGCAGCGGATG 4109885

RESULT 12  
US-08-581-918A-27  
Sequence 27, Application US/08581918A  
Patent No. 6043030  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Demetrick, Douglas J.  
APPLICANT: Serrano, Manuel  
APPLICANT: Hannan, Gregory J.  
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses  
TITLE OF INVENTION: Related Thereto  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordpad  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/581,918A  
FILING DATE: 02-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/497,214  
FILING DATE: 30-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/346,147  
FILING DATE: 29-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,511  
FILING DATE: 14-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,812  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,371  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/154,915  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,997  
FILING DATE: 17-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-071.06  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1299  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 401 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-581-918A-27

Query Match 78.9%; Score 15; DB 3; Length 401;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGTCTGCAGCGGGA 15  
DB 18 GGGTCTGCAGCGGGA 32

RESULT 13  
US-08-346-147B-27  
Sequence 27, Application US/08346147B  
Patent No. 6211334  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Demetrick, Douglas J.  
APPLICANT: Serrano, Manuel  
APPLICANT: Hannan, Gregory J.  
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses  
TITLE OF INVENTION: Related Thereto  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordpad  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/346,147B  
FILING DATE: 29-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,511  
FILING DATE: 14-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,812  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,371  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,915  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,997  
FILING DATE: 17-DEC-1992



ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-071.04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1299  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 401 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-346-147B-27

Query Match 78.9%; Score 15; DB 3; Length 401;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGA 15  
18 GGGTCTGCAGCGGA 32

Db

RESULT 14  
US-08-497-214D-27  
Sequence 27, Application US/08497214D  
Patent No. 6331390  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Demetrick, Douglas J.  
APPLICANT: Serrano, Manuel  
APPLICANT: Hannon, Gregory J.  
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordpad  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/497,214D  
FILING DATE: 30-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/346,147  
FILING DATE: 29-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,511  
FILING DATE: 14-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,812  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,371  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,915  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,997  
FILING DATE: 17-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-071.05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1299  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 401 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-497-214D-27

Query Match 78.9%; Score 15; DB 4; Length 401;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGA 15  
18 GGGTCTGCAGCGGA 32

Db

RESULT 15  
US-08-581-918A-26  
Sequence 26, Application US/08581918A  
Patent No. 6041030  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Demetrick, Douglas J.  
APPLICANT: Serrano, Manuel  
APPLICANT: Hannon, Gregory J.  
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordpad  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/581,918A  
FILING DATE: 02-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/497,214  
FILING DATE: 30-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/346,147  
FILING DATE: 29-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,511  
FILING DATE: 14-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,812  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,371  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,915  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,997  
FILING DATE: 17-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.

NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-071.06  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1299  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 404 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-581-918A-26

Query Match 78.9%; Score 15; DB 3; Length 404;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCTGCGAGCGGA 15  
DB 112 GGGCTGCGAGCGGA 126

RESULT 16  
US-08-346-147B-26  
Sequence 26, Application US/08346147B  
Patent No. 6211334  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Demetrick, Douglas J.  
APPLICANT: Serrano, Manuel  
APPLICANT: Hannon, Gregory J.  
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses  
TITLE OF INVENTION: Related Thereeto  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordpad  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/346,147B  
FILING DATE: 29-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,511  
FILING DATE: 14-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,812  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,371  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,915  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,997  
FILING DATE: 17-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-071.04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1299

TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 404 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-346-147B-26

Query Match 78.9%; Score 15; DB 3; Length 404;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCTGCGAGCGGA 15  
DB 112 GGGCTGCGAGCGGA 126

RESULT 17  
US-08-497-214D-26  
Sequence 26, Application US/08497214D  
Patent No. 6331390  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Demetrick, Douglas J.  
APPLICANT: Serrano, Manuel  
APPLICANT: Hannon, Gregory J.  
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses  
TITLE OF INVENTION: Related Thereeto  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordpad  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/497,214D  
FILING DATE: 30-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/346,147  
FILING DATE: 29-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,812  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,371  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,915  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,997  
FILING DATE: 17-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-071.05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1299  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:  
LENGTH: 404 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-497-214D-26

Query Match 78.9%; Score 15; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGTCTGCAGCGGA 15  
|||  
Db 112 GGGTCTGCAGCGGA 126

RESULT 18  
US-08-474-177-3  
Sequence 3, Application US/08474177  
Patent No. 5624819  
GENERAL INFORMATION:  
APPLICANT: Skolnick, Mark H.  
APPLICANT: Cannon-Albright, Lisa A.  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: GERMLINE MUTATIONS IN THE MTS GENE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,177  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03537  
FILING DATE: 17-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/251,938  
FILING DATE: 01-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348-E  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double

TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: Intron  
LOCATION: 1017..1149  
US-08-474-177-3

Query Match 78.9%; Score 15; DB 1; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGTCTGCAGCGGA 15  
|||  
Db 1026 GGGTCTGCAGCGGA 1040

RESULT 19  
US-08-467-033-3  
Sequence 3, Application US/08487033  
Patent No. 5739027  
GENERAL INFORMATION:  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: MTS1B1-Beta GENE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,033  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03316  
FILING DATE: 17-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348-C  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1017..1149  
US-08-487-033-3

Query Match 78.9%; Score 15; DB 1; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGGA 15  
|||||  
Db 1026 GGGTCTGCAGCGGGA 1040

RESULT 20  
US-08-480-810-3  
Sequence 3, Application US/08480810  
Patent No. 5801236  
GENERAL INFORMATION:  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: MTS1 GENE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,810  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03316  
FILING DATE: 17-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/251,938  
FILING DATE: 01-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1017..1149  
US-08-480-810-3

Query Match 78.9%; Score 15; DB 1; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGGA 15  
|||||  
Db 1026 GGGTCTGCAGCGGGA 1040

RESULT 21  
US-08-508-735-3  
Sequence 3, Application US/08508735  
Patent No. 5843756  
GENERAL INFORMATION:  
APPLICANT: Stone, Steven  
APPLICANT: Jiang, Ping  
TITLE OF INVENTION: MTS GENE AND THERAPEUTIC USE THEREOF  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/508,735  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US to be assigned  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03316  
FILING DATE: 17-MAR-1995  
ATTORNEY/AGENT INFORMATION:

NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4848  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: Intron  
LOCATION: 1017..1149  
US-08-508-735-3

Query Match 78.9%; Score 15; DB 2; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTCTGCGCGGA 15  
Db 1026 GGGTCTGCGCGGA 1040

RESULT 22  
US-08-848-251-3  
Sequence 3, Application US/08848251  
Patent No. 5989815  
GENERAL INFORMATION:  
APPLICANT: Skolnick, Mark H.  
APPLICANT: Cannon-Albright, Lisa A.  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: GERMLINE MUTATIONS IN THE MTS GENE AND  
TITLE OF INVENTION: METHOD FOR DETECTING PREDISPOSITION TO CANCER AT THE MTS  
NUMBER OF INVENTIONS: 36  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/848,251  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/474,083  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: PCT/US95/03537  
FILING DATE: 17-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/251,938  
FILING DATE: 01-JUN-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348-G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: Intron  
LOCATION: 1017..1149  
US-08-848-251-3

Query Match 78.9%; Score 15; DB 2; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTCTGCGCGGA 15  
Db 1026 GGGTCTGCGCGGA 1040

RESULT 23  
US-08-486-047-3  
Sequence 3, Application US/08486047  
Patent No. 5994095  
GENERAL INFORMATION:  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: MTS2 GENE  
NUMBER OF INVENTIONS: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,047  
FILING DATE: 07-JUN-1995

```
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/03316
FILING DATE: 17-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,938
FILING DATE: 01-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/215,087
FILING DATE: 18-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/215,086
FILING DATE: 18-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,369
FILING DATE: 14-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/214,582
FILING DATE: 18-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109348-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..866
NAME/KEY: CDS
LOCATION: 867..1016
FEATURE:
NAME/KEY: intron
LOCATION: 1017..1149
US-08-486-047-3

Query Match      78.9%; Score 15; DB 2; Length 1149;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGTCTGACGCGGA 15
DB      1026 GGGTCTGACGCGGA 1040

RESULT 24
US-09-120-130-3
; Sequence 3, Application US/09120130
; Patent No. 6037462
; GENERAL INFORMATION:
; APPLICANT: Kamb, Alexander
; TITLE OF INVENTION: MTSI GENE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
```

```
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/120,130
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,810
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,938
FILING DATE: 01-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/215,087
FILING DATE: 18-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/215,086
FILING DATE: 18-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,369
FILING DATE: 14-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/214,582
FILING DATE: 18-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109348
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..866
NAME/KEY: CDS
LOCATION: 867..1016
FEATURE:
NAME/KEY: intron
LOCATION: 1017..1149
US-09-120-130-3

Query Match      78.9%; Score 15; DB 3; Length 1149;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGTCTGACGCGGA 15
DB      1026 GGGTCTGACGCGGA 1040

RESULT 25
US-09-115-252-3
; Sequence 3, Application US/09115252
; Patent No. 6060301
; GENERAL INFORMATION:
; APPLICANT: Kamb, Alexander
; TITLE OF INVENTION: MTSI GENE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/115,252  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,810  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: PCT/US95/03316  
FILING DATE: 17-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/251,938  
FILING DATE: 01-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1017..1149  
US-09-115-252-3

Query Match 78.9%; Score 15; DB 3; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGA 15  
|||||  
DB 1026 GGGTCTGCAGCGGA 1040

RESULT 26  
US-08-986-515-3  
Sequence 3, Application US/08986515  
Patent No. 6090578  
GENERAL INFORMATION:  
APPLICANT: Kamd, Alexander  
TITLE OF INVENTION: MTS1 GENE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/986,515  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/480,810  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/251,938  
FILING DATE: 01-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1017..1149  
US-08-986-515-3

Query Match 78.9%; Score 15; DB 3; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGGTCTGCAGCGGA 15  
1026 GGGTCTGCAGCGGA 1040

RESULT 27  
US-09-120-128-3  
Sequence 3, Application US/09120128  
Patent No. 6140473  
GENERAL INFORMATION:  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: MTS2 GENE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/120,128  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,047  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: PCT/US95/03316  
FILING DATE: 17-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/251,938  
FILING DATE: 01-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:

NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: Intron  
LOCATION: 1017..1149  
US-09-120-128-3

Query Match 78.9%; Score 15; DB 3; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGA 15  
1026 GGGTCTGCAGCGGA 1040

RESULT 28  
US-09-120-129-3  
Sequence 3, Application US/09120129  
Patent No. 6180776  
GENERAL INFORMATION:  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: MTS2 GENE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/120,129  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/486,047  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: PCT/US95/03316  
FILING DATE: 17-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/251,938  
FILING DATE: 01-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear



MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: Intron  
LOCATION: 1017..1149  
US-09-120-129-3

Query Match 78.9%; Score 15; DB 3; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGACGCGGA 15  
|||  
Db 1026 GGGTCTGACGCGGA 1040

RESULT 29  
US-09-201-139-3  
Sequence 3, Application US/09201139  
Patent No. 6210949  
GENERAL INFORMATION:  
APPLICANT: Stone, Steven  
APPLICANT: Jiang, Ping  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: MTS GENE AND THERAPEUTIC USE THEREOF  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/201.139  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/508,735  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03316  
FILING DATE: 17-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4848  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: Intron  
LOCATION: 1017..1149  
US-09-201-139-3

Query Match 78.9%; Score 15; DB 3; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGACGCGGA 15  
|||  
Db 1026 GGGTCTGACGCGGA 1040

RESULT 30  
US-09-120-131-3  
Sequence 3, Application US/09120131  
Patent No. 6218146  
GENERAL INFORMATION:  
APPLICANT: Kamb, Alexander  
TITLE OF INVENTION: MTS2 GENE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/120.131  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/486,047  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: PCT/US95/03316  
FILING DATE: 17-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/251,938  
FILING DATE: 01-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,087  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,086  
FILING DATE: 18-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,369  
FILING DATE: 14-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/214,582  
FILING DATE: 18-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109348-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1149 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..866  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 867..1016  
FEATURE:  
NAME/KEY: Intron  
LOCATION: 1017..1149  
US-09-120-131-3

Query Match 78.9%; Score 15; DB 3; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCTGCGCGGGA 15  
DB 1026 GGGTCTGCGCGGGA 1040

RESULT 31  
US-08-311-731A-130  
Sequence 130, Application US/08311731A  
Patent No. 6583286  
GENERAL INFORMATION:  
APPLICANT: SMITH, DOUGLAS  
APPLICANT: MAO, JEN-I  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LARPAE FOR  
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 411  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
STREET: 600 ATLANTIC AVENUE  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,731A  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: GATES, EDWARD R.  
REGISTRATION NUMBER: 31,616  
REFERENCE/DOCKET NUMBER: C0044/7125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/720-3500  
TELEFAX: 617/720-2441  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36941 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO

ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: MYCOBACTERIUM LEPRAE  
US-08-311-731A-130

Query Match 78.9%; Score 15; DB 4; Length 36941;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CTGCAGCGCGGATGCT 19  
DB 12443 CTGCAGCGCGGATGCT 12457

RESULT 32  
US-09-595-386-11  
Sequence 11, Application US/09595386  
Patent No. 6544792  
GENERAL INFORMATION:  
APPLICANT: Diaz-Torres, Maria  
TITLE OF INVENTION: Production of Secreted Polypeptides  
FILE REFERENCE: GC618  
CURRENT APPLICATION NUMBER: US/09/595,386  
CURRENT FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: WO PCT/US99/31010  
PRIOR FILING DATE: 1999-12-21  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 52  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: primer  
US-09-595-386-11

Query Match 77.9%; Score 14.8; DB 4; Length 52;  
Best Local Similarity 88.9%; Pred. No. 1.9e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCGCGGATG 18  
DB 3 GCGCTGCGCGGATG 20

RESULT 33  
US-08-750-088A-69/C  
Sequence 69, Application US/08750088A  
Patent No. 6329138  
GENERAL INFORMATION:  
APPLICANT: DE BRENNOWER, HANS  
APPLICANT: PORTAELS, FRAN OISE  
APPLICANT: WACHTELINCKX, LIEVE  
APPLICANT: JANNES, GEBERT  
APPLICANT: ROSSAU, RUDI  
TITLE OF INVENTION: METHOD FOR DETECTION OF THE ANTIBIOTIC  
TITLE OF INVENTION: RESISTANCE SPECTRUM OF MYCOBACTERIUM SPECIES  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: US  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,088A  
FILING DATE: 21-FEB-1997

```

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1657.0010000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-750-088A-69

Query Match 77.9%; Score 14.8; DB 4; Length 228;
Best Local Similarity 88.9%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGG 18
Db 76 GGGTCTGCAGCGGATGG 59

RESULT 34
US-09-147-935A-5/C
; Sequence 5, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 5
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium avium
; US-09-147-935A-5

Query Match 77.9%; Score 14.8; DB 3; Length 306;
Best Local Similarity 88.9%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGG 18
Db 117 GGGTCTGCAGCGGATGG 100

RESULT 35
US-09-147-935A-11/C
; Sequence 11, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
```

```

; SEQ ID NO 11
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium chitae
; US-09-147-935A-11

Query Match 77.9%; Score 14.8; DB 3; Length 306;
Best Local Similarity 88.9%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGG 18
Db 117 GGGTCTGCAGCGGATGG 100

RESULT 36
US-09-147-935A-17/C
; Sequence 17, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 17
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium genavense
; US-09-147-935A-17

Query Match 77.9%; Score 14.8; DB 3; Length 306;
Best Local Similarity 88.9%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGG 18
Db 117 GGGTCTGCAGCGGATGG 100

RESULT 37
US-09-147-935A-18/C
; Sequence 18, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 18
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium goodii
; US-09-147-935A-18

Query Match 77.9%; Score 14.8; DB 3; Length 306;
Best Local Similarity 88.9%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGG 18
```

Db 117 GGGTCTGCGCGGTATGG 100

RESULT 38  
US-09-147-935A-22/c  
; Sequence 22, Application US/09147935A  
; Patent No. 6242584  
; GENERAL INFORMATION:  
; APPLICANT: KOOK, Yoon-Hoh  
; APPLICANT: KIM, Bum-Joon  
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY  
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE  
; FILE REFERENCE: 0136/0F425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/KR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 22  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium intracellulare  
US-09-147-935A-22

Query Match 77.9%; Score 14.8; DB 3; Length 306;  
Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCGCGGTATGG 18  
Db 117 GGGTCTGCGCGGTATGG 100

RESULT 39  
US-09-147-935A-28/c  
; Sequence 28, Application US/09147935A  
; Patent No. 6242584  
; GENERAL INFORMATION:  
; APPLICANT: KOOK, Yoon-Hoh  
; APPLICANT: KIM, Bum-Joon  
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY  
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE  
; FILE REFERENCE: 0136/0F425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/KR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 28  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium nonchromogenicum  
US-09-147-935A-28

Query Match 77.9%; Score 14.8; DB 3; Length 306;  
Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCGCGGTATGG 18  
Db 117 GGGTCTGCGCGGTATGG 100

RESULT 40  
US-09-147-935A-29/c  
; Sequence 29, Application US/09147935A  
; Patent No. 6242584  
; GENERAL INFORMATION:  
; APPLICANT: KOOK, Yoon-Hoh  
; APPLICANT: KIM, Bum-Joon

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY  
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE  
; FILE REFERENCE: 0136/0F425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/KR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 29  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium paratuberculosis  
US-09-147-935A-29

Query Match 77.9%; Score 14.8; DB 3; Length 306;  
Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGTCTGCGCGGTATGG 18  
Db 117 GGGTCTGCGCGGTATGG 100

Search completed: January 13, 2004, 17:58:34  
Job time : 49.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 13, 2004, 15:53:18 / Search time 744 Seconds  
(without alignment)  
1044.735 Million cell updates/sec

Title: US-09-719-737-9

Perfect score: 19  
Sequence: 1 gggctcgcagcgagatgct 19

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmb1:  
1: gb\_ba:\*  
2: gb\_bhg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vtc:\*  
38: em\_hvg:\*  
39: em\_hvg\_hum:\*  
40: em\_hvg\_mus:\*  
41: em\_hvg\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	19	100.0	19	6	AX008655	AX008655 Sequence
2	19	100.0	19	6	AX008656	AX008656 Sequence
3	19	100.0	19	6	AX671168	AX671168 Sequence
4	19	100.0	19	6	AX671179	AX671179 Sequence
5	19	100.0	19	6	AX671180	AX671180 Sequence
6	19	100.0	2996	9	HUMGCMSPFB	MS9941 Human GM-CS
7	19	100.0	86574	9	HS83387	AL008637 Human DNA
8	18	94.7	19	6	AX008657	AX008657 Sequence
9	18	94.7	19	6	AX008658	AX008658 Sequence
10	18	94.7	19	6	AX671169	AX671169 Sequence
11	18	94.7	19	6	AX671170	AX671170 Sequence
12	17.4	91.6	783	11	BV043795	BV043795 S212P6415
13	17.4	91.6	2711	10	CPU94688	U94688 Cavia porce
14	17.4	91.6	67292	2	AC120698_3	Continuation (4 of
15	17.4	91.6	86257	2	AL954650	AL954650 Human DNA
16	17.4	91.6	110000	2	AC120698_2	Continuation (3 of
17	17.4	91.6	142509	2	AC117942	AC117942 Tetradon
18	17.4	91.6	164532	2	AC116440	AC116440 Tetradon
19	17.4	91.6	177816	2	AC105219	AC105219 Homo sapi
20	17.4	91.6	200315	2	AC104885	AC104885 Mus muscu
21	17.4	91.6	246937	2	AC111254	AC111254 Rattus no
22	17	89.5	19	6	AX008662	AX008662 Sequence
23	17	89.5	19	6	AX671174	AX671174 Sequence
24	17	89.5	19	6	AX671181	AX671181 Sequence
25	17	89.5	19	6	AX671182	AX671182 Sequence
26	17	89.5	10225	1	AE004803	AE004803 Pseudomon
27	16.4	86.3	436	6	AX066133	AX066133 Sequence
28	16.4	86.3	778	9	AB02081540	AB020815 Homo sapi
29	16.4	86.3	1468	1	AB006797	AB006797 Pseudomon
30	16.4	86.3	1824	6	AX120193	AX120193 Sequence
31	16.4	86.3	1824	6	BD162310	BD162310 Novel pol
32	16.4	86.3	1947	6	AX066131	AX066131 Sequence
33	16.4	86.3	2097	9	BC015649	BC015649 Homo sapi
34	16.4	86.3	4690	1	MLCPMOA	LA0804 Methylcocce
35	16.4	86.3	5802	6	A90822	A90822 Sequence 1
36	16.4	86.3	5802	6	AR193170	AR193170 Sequence
37	16.4	86.3	5802	6	BD064963	BD064963 Protein c
38	16.4	86.3	5802	9	HSDMBT1	AI000342 Homo sapi
39	16.4	86.3	7080	1	MCU94337	U94337 Methylcocce
40	16.4	86.3	7656	9	HS243224	AJ243224 Homo sapi
41	16.4	86.3	7683	9	HS243212	AJ243212 Homo sapi
42	16.4	86.3	7686	1	AF159456	AF159456 Homo sapi
43	16.4	86.3	10171	1	AE004704	AE004704 Pseudomon
44	16.4	86.3	10698	1	AE005902	AE005902 Caulobact
45	16.4	86.3	12239	1	AE012048	AE012048 Xanthomon

ALIGNMENTS

RESULT 1  
AX008655/c  
LOCUS AX008655 19 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 8 from Patent WO9966037.  
ACCESSION AX008655  
VERSION AX008655.1 GI:9996179  
KEYWORDS  
ORGANISM  
SOURCE  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1  
AUTHORS  
TITLE  
JOURNAL  
Renzl, P.  
Antisense oligonucleotides for treating or preventing atopic  
diseases and neoplastic cell proliferation  
Patent: WO 9966037-A 8 23-DEC-1999;

RENTZI PAOLO (CA) ; RECH EXPERTISES ET DEV MEDICAU (CA)  
Location/Qualifiers

FEATURES  
source  
1. 19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Sense oligonucleotide for IL-3, IL-5 and GM-CSF"

BASE COUNT  
4 a 10 c 3 g 2 t

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 19; DB 6; Length 19;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  
1 GGGTCTGCAGCGGATGCT 19  
19 GGGTCTGCAGCGGATGCT 1

Db  
19 GGGTCTGCAGCGGATGCT 1

RESULT 2  
AX008656 19 bp DNA linear PAT 06-SEP-2000  
LOCUS  
DEFINITION Sequence 9 from Patent WO9966037.  
ACCESSION AX008656  
VERSION AX008656.1 GI:9996180  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.

REFERENCE  
1  
AUTHORS  
TITLE  
JOURNAL  
RENTZI P.  
Antisense oligonucleotides for treating or preventing atopic diseases and neoplastic cell proliferation  
Patent: WO 9966037-A 9 23-DEC-1999;  
RENTZI PAOLO (CA) ; RECH EXPERTISES ET DEV MEDICAU (CA)  
Location/Qualifiers

FEATURES  
source  
1. 19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Antisense oligonucleotide inhibiting the common subunit of IL-3, IL-5 and GM-CSF human receptor"

BASE COUNT  
2 a 3 c 10 g 4 t

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 19; DB 6; Length 19;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  
1 GGGTCTGCAGCGGATGCT 19  
1 GGGTCTGCAGCGGATGCT 19

Db  
1 GGGTCTGCAGCGGATGCT 19

RESULT 3  
AX671168 19 bp DNA linear PAT 27-MAR-2003  
LOCUS  
DEFINITION Sequence 8 from Patent WO03004511.  
ACCESSION AX671168  
VERSION AX671168.1 GI:29329624  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.

REFERENCE  
1  
AUTHORS  
TITLE  
JOURNAL  
RENTZI P., Allam M. and Allakhverdi Z.  
Methods for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
Patent: WO 03004511-A 8 16-JAN-2003;  
Topigen Pharmaceuticals Inc (CA)  
Location/Qualifiers

FEATURES  
source  
1. 19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Sequence is completely synthesized"

/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Sequence is completely synthesized"

BASE COUNT  
2 a 3 c 10 g 4 t

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 19; DB 6; Length 19;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  
1 GGGTCTGCAGCGGATGCT 19  
1 GGGTCTGCAGCGGATGCT 19

Db  
1 GGGTCTGCAGCGGATGCT 19

RESULT 4  
AX671179/c 19 bp DNA linear PAT 27-MAR-2003  
LOCUS  
DEFINITION Sequence 19 from Patent WO03004511.  
ACCESSION AX671179  
VERSION AX671179.1 GI:29329635  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.

REFERENCE  
1  
AUTHORS  
TITLE  
JOURNAL  
RENTZI P., Allam M. and Allakhverdi Z.  
Methods for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
Patent: WO 03004511-A 19 16-JAN-2003;  
Topigen Pharmaceuticals Inc (CA)  
Location/Qualifiers

FEATURES  
source  
1. 19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Sequence is completely synthesized"

BASE COUNT  
4 a 10 c 3 g 2 t

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 19; DB 6; Length 19;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  
1 GGGTCTGCAGCGGATGCT 19  
19 GGGTCTGCAGCGGATGCT 1

Db  
19 GGGTCTGCAGCGGATGCT 1

RESULT 5  
AX671180 19 bp DNA linear PAT 27-MAR-2003  
LOCUS  
DEFINITION Sequence 20 from Patent WO03004511.  
ACCESSION AX671180  
VERSION AX671180.1 GI:29329636  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.

REFERENCE  
1  
AUTHORS  
TITLE  
JOURNAL  
RENTZI P., Allam M. and Allakhverdi Z.  
Methods for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
Patent: WO 03004511-A 20 16-JAN-2003;  
Topigen Pharmaceuticals Inc (CA)  
Location/Qualifiers

FEATURES  
source  
1. 19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Sequence is completely synthesized"

Query Match 100.0%; Score 19; DB 6; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTCTGACGGGATGCT 19  
 |||||  
 1 GGGTCTGACGGGATGCT 19

RESULT 6  
 LOCUS H0MGCSPFB 2996 bp mRNA linear PRI 16-MAY-1994  
 DEFINITION Human GM-CSF receptor beta chain mRNA, complete cds.  
 ACCESSION M59941 M38275  
 VERSION M59941.1 GI:487424  
 KEYWORDS GM-CSF receptor; cytokine receptor; growth factor receptor;  
 lymphokine receptor;  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2996)  
 Hayashida, K., Kitamura, T., Gorman, D.M., Arai, K., Yokota, T. and  
 Miyajima, A.  
 Molecular cloning of a second subunit of the receptor for human  
 granulocyte-macrophage colony-stimulating factor (GM-CSF):  
 reconstitution of a high-affinity GM-CSF receptor  
 Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9655-9659 (1990)  
 91088571  
 PUBMED 1702217  
 2 (bases 1 to 2996)  
 Kitamura, T.  
 Direct Submission  
 TITLE Submitted (05-FEB-1991) Toshio Kitamura, Department of Molecular  
 Biology, DNA Research Institute of Molecular and Cellular Biology,  
 901 California Avenue, Palo Alto, CA 94304-1104, USA  
 On May 16, 1994 this sequence version replaced gi:183367.  
 COMMENT Original source text: Human CDNA to mRNA.  
 FEATURES  
 source location/Qualifiers  
 1..2996  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 29..2722  
 /codon\_start=1  
 /product="GM-CSF receptor beta chain"  
 /protein\_id="AA18171.1"  
 /db\_xref="GI:487425"  
 /translation="MVLGGGLSMALALCWERSLAGAERTPLQTLRCYNDYTSHT  
 CRNADODAGRLVNVTLIRRVNEDLEPVSGLSDMDMPCACHPRCVPRCVPCOS  
 FVTVDDYFQPRPLGRTUTLTTHVQPPERDQITDQHTLITLTSVALGSGO  
 SHMLFDELFVYVYKRLQSDMEDAALLNSTQATGPHLMPSSTYVARVRLAP  
 GSRISGRSPKSPVCEWDSQDEGDAVCEPDAVSCSWKVEVASVSEGL  
 FYRSPDAGEECSPVIREGLSLTHHCQIPVDPATGQIVSVQPRAEKGLS  
 SVNIQMPRLNTVTKDGSYSLRWETMKMYEHIDHTEFIOYKDTATWDSKETLO  
 NAIMSMLPALEPSTRVAVRVRSTRGNGSESEARSMDTESYLPMWALATYI  
 FLITAVILARFCGTYRARRKREKIPNSKSHLFQNSAELIMPQMSATPSGP  
 PHQPMGSRFPPELGVFPVGGDESEPLTIEDKHVCDPSPGDTTPASDLPTSE  
 PPSQPGPASHSTPEKQASSPFENGPGYLPSPHSRLDILGQPPQEGSOKSPPG  
 SLEVLCLPAGGVOVLPLAQMGPGQAVEVERSPQAGASPISLESQGPAPALGPR  
 VGGODQDSVPAIEMSGDTEDEPGVSSADLVTPNSGASVSLVSLGLPSPO  
 TPLCLPGLASGPGAPGPGVSGFEGYVELPPIEGRSRSPRNNVPEAKSPVNPCE  
 RPDVSTSPQPEGLLVLYQVGVCTPLPGIGPGLSLSRKSPSSGPGPEPEIKNDQAO  
 VKRPGQAVPQVPTQLFKALKQDDYSLPPEVNRKGEVC"  
 29..172  
 173..2719  
 /product="GM-CSF receptor beta chain"  
 /db\_xref="taxon:9606"  
 591 a 1017 c 857 g 531 t  
 ORIGIN  
 BASE COUNT 591 a 1017 c 857 g 531 t  
 Query Match 100.0%; Score 19; DB 9; Length 2996;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GGGTCTGACGGGATGCT 19  
 |||||  
 125 GGGTCTGACGGGATGCT 107

RESULT 7  
 LOCUS HS833B7 86574 bp DNA linear PRI 05-JUN-2003  
 DEFINITION Human DNA sequence from clone CTA-833B7 on chromosome 22q12.3-13.2  
 Contains the NCF4 gene for cytosolic neutrophil factor 4 (40kd),  
 the 5' part of the CSF2RB gene for granulocyte-macrophage  
 low-affinity colony stimulating factor 2 receptor beta, ESF6, STS  
 and GSS6, complete sequence.  
 ACCESSION AL008637  
 VERSION AL008637.1 GI:313600  
 KEYWORDS HTG; colony stimulating factor; CSF2RB; cytosolic neutrophil  
 factor; NCF4.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 86574)  
 Burton, J.  
 Direct Submission  
 TITLE Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk  
 On May 18, 1998 this sequence version replaced gi:2578146.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information  
 on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep CTA-833B7 is  
 from the human BAC library described in U-J. Kim et al. (1996)  
 Genomics 34, 213-218.  
 VECTOR: pBelBAC11

FEATURES  
 source location/Qualifiers  
 1..86574  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="22"  
 /map="q12.3-13.2"

Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquerry@sanger.ac.uk  
 -----  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest, except on the rare  
 occasion of the clone being a YAC.  
 This sequence was generated from part of bacterial clone contigs of  
 human chromosome 22, constructed by the Sanger Centre Chromosome 22  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr22  
 IMPORTANT: This sequence is not the entire insert of clone  
 CTA-833B7 it may be shorter because we sequence overlapping  
 sections only once, except for a short overlap.  
 The true right end of clone CTFP22-24B5 is at 100 in this sequence.  
 location/Qualifiers  
 1..86574  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="22"  
 /map="q12.3-13.2"

```

/clone="CTA-833B7"
/clone_lib="CT978SK-A2"
repeat_region
  /note="AluSp repeat: matches 138. .313 of consensus"
  1. .176
repeat_region
  /note="2.0 copies 5 mer GAAA 20% conserved"
  211. .220
repeat_region
  /note="2.5 copies 6 mer CCGCC 30% conserved"
  235. .249
repeat_region
  /note="2.5 copies 1 mer C 22% conserved"
  239. .258
repeat_region
  /note="2.0 copies 1 mer C 22% conserved"
  627. .636
repeat_region
  /note="2.5 copies 4 mer TCCC 20% conserved"
  642. .653
repeat_region
  /note="4.0 copies 3 mer AGA 24% conserved"
  complement(862. .903)
repeat_region
  /note="MIR repeat: matches 212. .254 of consensus"
  1072. .1084
repeat_region
  /note="2.6 copies 5 mer GGAGA 26% conserved"
  1163. .1173
repeat_region
  /note="2.2 copies 5 mer GGAGT 22% conserved"
  complement(1233. .1462)
repeat_region
  /note="L2 repeat: matches 2418. .2653 of consensus"
  1791. .2095
repeat_region
  /note="AluX repeat: matches 1. .307 of consensus"
  2521. .2537
repeat_region
  /note="4.2 copies 4 mer CTC 25% conserved"
  2866. .2881
repeat_region
  /note="2.0 copies 8 mer GAGGAG 32% conserved"
  2902. .2911
repeat_region
  /note="2.5 copies 4 mer GGAG 20% conserved"
  3263. .3277
repeat_region
  /note="2.1 copies 7 mer CCTGACC 30% conserved"
  3359. .3522
repeat_region
  /note="MERS8A repeat: matches 43. .207 of consensus"
  3532. .3838
repeat_region
  /note="AluX repeat: matches 1. .308 of consensus"
  complement(3872. .3956)
repeat_region
  /note="MLTJ2 repeat: matches 366. .450 of consensus"
  complement(4030. .4288)
repeat_region
  /note="MLTJ repeat: matches 1. .261 of consensus"
  complement(4478. .4555)
repeat_region
  /note="MIR repeat: matches 64. .143 of consensus"
  4592. .4609
repeat_region
  /note="3.0 copies 6 mer CTGGG 27% conserved"
  complement(4645. .5195)
repeat_region
  /note="LTR50 repeat: matches 4. .766 of consensus"
  5291. .5370
repeat_region
  /note="LIME1 repeat: matches 5523. .5604 of consensus"
  complement(5371. .5676)
repeat_region
  /note="AluX repeat: matches 1. .310 of consensus"
  5677. .6268
repeat_region
  /note="LIME1 repeat: matches 5604. .6332 of consensus"
  6302. .6481
repeat_region
  /note="L2 repeat: matches 2636. .2819 of consensus"
  complement(6483. .6755)
repeat_region
  /note="MER115 repeat: matches 393. .688 of consensus"
  6934. .6962
repeat_region
  /note="2.1 copies 14 mer TGCATGATGTGA 49% conserved"
  6944. .6953
repeat_region
  /note="2.5 copies 4 mer TGCA 20% conserved"
  6984. .7002
repeat_region
  /note="1.9 copies 10 mer TGCAAGGTG 29% conserved"
  6999. .7085
repeat_region
  /note="6.2 copies 14 mer GTGTATGTACT 66% conserved"
  7000. .7014
repeat_region
  /note="2.5 copies 6 mer TGTACA 30% conserved"
  7006. .7022
repeat_region
  /note="2.1 copies 8 mer TGTACATG 25% conserved"
  7047. .7193
repeat_region
  /note="6.1 copies 24 mer GTGTGATGTGTATGTGCTTCG 113% conserved"
  7048. .7080
/note="5.5 copies 6 mer TGTGTA 30% conserved"

```

```

repeat_region
  7053. .7085
repeat_region
  /note="4.1 copies 8 mer ATGTGTAT 29% conserved"
  7118. .7127
repeat_region
  /note="2.5 copies 4 mer TGCA 20% conserved"
  7125. .7136
repeat_region
  /note="2.0 copies 6 mer ATGTGT 24% conserved"
  7148. .7164
repeat_region
  /note="2.8 copies 6 mer TGTGTA 25% conserved"
  7203. .7332
repeat_region
  /note="L2 repeat: matches 3165. .3313 of consensus"
  7333. .7342
repeat_region
  /note="2.5 copies 4 mer GAAT 20% conserved"
  7347. .7532
repeat_region
  /note="MIR repeat: matches 2. .191 of consensus"
  complement(7537. .7640)
repeat_region
  /note="MIR repeat: matches 158. .262 of consensus"
  complement(8152. .8228)
repeat_region
  /note="MIR repeat: matches 184. .260 of consensus"
  8253. .8341
repeat_region
  /note="L2 repeat: matches 3214. .3311 of consensus"
  8504. .8516
repeat_region
  /note="2.6 copies 5 mer AGAGA 26% conserved"
  8848. .8859
repeat_region
  /note="2.0 copies 6 mer AGTTC 24% conserved"
  9588. .9602
repeat_region
  /note="5.0 copies 3 mer GCA 21% conserved"
  9920. .9935
repeat_region
  /note="2.0 copies 8 mer TTGGGGG 32% conserved"
  10303. .10313
repeat_region
  /note="2.2 copies 5 mer ACACA 22% conserved"
  10428. .10439
repeat_region
  /note="3.0 copies 4 mer TGCT 24% conserved"
  10570. .10580
repeat_region
  /note="2.2 copies 5 mer TTTC 22% conserved"
  complement(11386. .11850)
repeat_region
  /note="L2 repeat: matches 2722. .3312 of consensus"
  11913. .29288
gene
  /gene="NCP4"
  join(11913. .12573,15415. .15499,16289. .16442,18762. .18832,
  21785. .21912,23022. .23079,23697. .23795,27023. .27153,
  27359. .27464,28998. .29288)
  /gene="NCP4"
  /product="pK833B7.1 (neutrophil cytosolic factor 4
  (40kD))"
  /note="match: CDNA8: Em:AB002665 Em:X77094 Em:U59488
  match: ESTs: Em:AA177839 Em:A1007048 Em:AA465462
  Em:AA465389 Em:AA485518 Em:A181940 Em:AA744805
  Em:AA975113 Em:AA969460 Em:AA984420 Em:A1435296
  Em:A1299103 Em:AA648472 Em:AA702857 Em:A1088359 Em:DD0144
  Em:AA909156 Em:AA688071 Em:W95229 Em:A1439568"
  /evidence=not experimental
  complement(11582. .12109)
repeat_region
  /note="L2 repeat: matches 2918. .3064 of consensus"
  12358. .12367
repeat_region
  /note="3.3 copies 3 mer GGA 20% conserved"
  12443. .12453
repeat_region
  /note="2.2 copies 5 mer GAGAC 22% conserved"
  join(12542. .12573,15415. .15499,16289. .16442,18762. .18832,
  21785. .21912,23022. .23079,23697. .23795,27023. .27153,
  27359. .27464,28998. .29193)
  /gene="NCP4"
  /note="match: proteins: Tr:O60808 Tr:P97369"
  /codon_start=1

```

```

Query Match      100.0%; Score 19; DB 9; Length 86574;
Beat Local Similarity 100.0%; Pred. No. 67;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GGGTGTGAGGGGAGTGT 19
DB      74634 GGGTGTGAGGGGAGTGT 74616

```



RESULT 8  
LOCUS AX008657 19 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 10 from Patent WO9966037.  
ACCESSION AX008657  
VERSION AX008657.1 GI:9996181  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Renzi, P.  
TITLE Antisense oligonucleotides for treating or preventing atopic diseases and neoplastic cell proliferation  
JOURNAL Patent: WO 9966037-A 10 23-DEC-1999;  
RENZI PAOLO (CA); RECH EXPERTISES ET DEV MEDICAU (CA)  
FEATURES  
Source  
1. .19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Antisense oligonucleotide inhibiting the common subunit of IL-3, IL-5 and GM-CSF human receptor"  
BASE COUNT 2 a 3 c 9 g 5 t  
ORIGIN  
Query Match 94.7%; Score 18; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GGCTGCAGCGGATGCT 19  
Db 1 GGCTGCAGCGGATGCT 18  
RESULT 9  
LOCUS AX008658 19 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 11 from Patent WO9966037.  
ACCESSION AX008658  
VERSION AX008658.1 GI:9996182  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Renzi, P.  
TITLE Antisense oligonucleotides for treating or preventing atopic diseases and neoplastic cell proliferation  
JOURNAL Patent: WO 9966037-A 11 23-DEC-1999;  
RENZI PAOLO (CA); RECH EXPERTISES ET DEV MEDICAU (CA)  
FEATURES  
Source  
1. .19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Antisense oligonucleotide inhibiting the common subunit of IL-3, IL-5 and GM-CSF human receptor"  
BASE COUNT 3 a 3 c 10 g 3 t  
ORIGIN  
Query Match 94.7%; Score 18; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGCTCTGCAGCGGATGG 18  
Db 2 GGCTCTGCAGCGGATGG 19  
RESULT 10  
LOCUS AX671169 19 bp DNA linear PAT 27-MAR-2003

DEFINITION Sequence 9 from Patent WO03004511.  
ACCESSION AX671169  
VERSION AX671169.1 GI:29329625  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Renzi, P., Allam, M. and Allakhverdi, Z.  
TITLE Methods for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
JOURNAL Patent: WO 03004511-A 9 16-JAN-2003;  
Topigen Pharmaceuticals Inc (CA)  
FEATURES  
Source  
1. .19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Sequence is completely synthesized"  
BASE COUNT 2 a 3 c 9 g 5 t  
ORIGIN  
Query Match 94.7%; Score 18; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GGCTGCAGCGGATGCT 19  
Db 1 GGCTGCAGCGGATGCT 18  
RESULT 11  
LOCUS AX671170 19 bp DNA linear PAT 27-MAR-2003  
DEFINITION Sequence 10 from Patent WO03004511.  
ACCESSION AX671170  
VERSION AX671170.1 GI:29329626  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Renzi, P., Allam, M. and Allakhverdi, Z.  
TITLE Methods for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
JOURNAL Patent: WO 03004511-A 10 16-JAN-2003;  
Topigen Pharmaceuticals Inc (CA)  
FEATURES  
Source  
1. .19  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Sequence is completely synthesized"  
BASE COUNT 3 a 3 c 10 g 3 t  
ORIGIN  
Query Match 94.7%; Score 18; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGCTCTGCAGCGGATGG 18  
Db 2 GGCTCTGCAGCGGATGG 19  
RESULT 12  
LOCUS BV043795/c 783 bp DNA linear STS 31-MAY-2003  
DEFINITION S212P6415FAL0.T0 CZECHII/EI Mus musculus STS genomic, sequence tagged site.  
ACCESSION BV043795  
VERSION BV043795.1 GI:31159590  
KEYWORDS STS.

SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 783)  
AUTHORS Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C., Lander, E.S., Lindblad-Toh, K. and Daly, M.J.  
TITLE The mosaic structure of variation in the laboratory mouse genome  
JOURNAL Nature 420 (6915), 574-578 (2002)  
MEDLINE 22354684  
PUBMED 12468852

COMMENT Contact: Kerstin Lindblad-Toh  
Whitehead Institute for Biomedical Research, Center for Genome Research  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172521477  
Fax: 6172580903  
Email: kersti@genome.wi.mit.edu  
Primer A: None  
Primer B: None  
STS size: 783  
Protocol:  
WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 12951/SVIMJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGSv3 C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

FEATURES  
source 1..783  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/map="16 40-713 30321911-30322586"  
/clone\_id="C57BL/6J"  
Location/Qualifiers

BASE COUNT 180 a 202 c 199 g 201 t 1 others  
ORIGIN

Query Match 91.6%; Score 17.4; DB 11; Length 783;  
Best Local Similarity 94.7%; Pred. No. 7.5e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||  
568 GGGCTCTGCAGCGGATGCT 550

RESULT 13  
CPU94688 2711 bp mRNA linear ROD 16-NOV-1998  
LOCUS CPU94688/c  
DEFINITION Cavia porcellus interleukin-5 receptor beta chain (IL5) mRNA, complete cds.  
ACCESSION U94688  
VERSION U94688.1 GI:3882357  
KEYWORDS  
SOURCE Cavia porcellus (domestic guinea pig)  
ORGANISM Cavia porcellus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
REFERENCE 1 (bases 1 to 2711)  
AUTHORS Logsdon, N.J., Graham, A. and Scott, C.W.  
TITLE Guinea pig IL5 receptor beta chain  
JOURNAL Unpublished  
2 (bases 1 to 2711)  
Logsdon, N.J., Graham, A. and Scott, C.W.  
Direct Submission  
Submitted (19-MAR-1997) Target Discovery, Zeneca Pharmaceuticals,  
1800 Concord Pike, Wilmington, DE 19850-5437, USA

FEATURES  
source Location/Qualifiers  
1..2711  
/organism="Cavia porcellus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10141"  
/cell\_type="bronchial alveolar lavage cells"  
gene 1..2711  
/gene="IL5"  
17..2689  
/gene="IL5"  
/codon\_start=1  
/product="interleukin-5 receptor beta chain"  
/protein\_id="AAC77520.1"  
/db\_xref="GI:3882358"  
translation="MAPPGLLMALLLLCRPPVEGSEDTVPQLTRCYNDYTSRIYCSMAEABAEOLIVTLHRRRLKSNHSEAVSCPELTDMMSHCSPSCVPRCVIPYTRALDNDYSPEDPRPLDRLTYLTAHVOPPPDDVINTSGDVLITWAVALEGPHTSWSQRLREVEYTKLHEHWEASSTHNSSQALAPFLPFBSTYVAVVRRLARGSGRPSQMSPEVSWSSQPDQKQPNLQCFPGAHITLCSMWRNCOVTSVSFGLPYRSILDAGEQECQVQKELHDYTRSCQIRVSNRPHSQYTVVPRNGEKPIRSANHIOMAPLTNAVTKDQDYSILRWTEKMYSHIENTPEIQRTAGDMWENKXTLNKHNMPLEPLPATYTLARVRKSPGAGNGIMSESEORMTDMLPTWVLA LVLYLTALLALRFGCLXGYRLNRKKEKIPNPSKSHLPKNSAGLRPLDSMAFALSAPSWCMGKRFLEIGVCPADSRSEVPLITTEPTVCPDPSERHSPASDLTQEPSPVQPGPPVQDDQDLATFDENGPIGPHSHSLPDLAQQKPELPSGLSLYLCLEPGRALQVLPAQATQAQAPAGCISGVPYQGSPLYLACGGAPAPPSDPSQAQGGPGGPDSPVILPTNSGGEHVPVAGVGYTTDALITLSTEASVSLAPPDLCGQSDRAPPAPTPGKPCFGLPSPAGMGLPFGGFPVAPAPSPVLSPTGQDVP LSPAPBGLVTCQVGDGCEFLPGLGSGFLGSGSPSKSPSPVPCPEINDIEGFFVKKPPGQ PMPQVPAIQFKSLKQDDYTLTPMWRSPRPHVC"

BASE COUNT 513 a 940 c 777 g 481 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 2711;  
Best Local Similarity 94.7%; Pred. No. 6.4e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||  
Db 113 GGGTCTGCAGCGGATGCT 95

RESULT 14  
AC120698 3/c  
LOCUS AC120698 Accession AC120698  
WCOMMENT  
Sequence split into 4 fragments  
Fragment Name Begin End  
AC120698\_0 1 110000  
AC120698\_1 100001 210000  
AC120698\_2 200001 310000  
AC120698\_3 300001 367292  
Continuation (4 of 4) of AC120698 from base 300001 (AC120698 Rattus norvegicus clone CH2;

Query Match 91.6%; Score 17.4; DB 2; Length 67292;  
Best Local Similarity 94.7%; Pred. No. 4.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCAGCGGATGCT 19  
|||  
Db 2383 GGGTCTGCAGCGGATGCT 2365

RESULT 15  
AL954650 86257 bp DNA linear PRI 26-FEB-2003  
LOCUS AL954650/c  
DEFINITION Human DNA sequence from clone Rpl1-533M19 on chromosome 1, complete sequence.  
ACCESSION AL954650  
VERSION AL954650.8 GI:28564367  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Mammalia; Eutheria; Primates; Carnivora; Hominiidae; Homo.  
1 (bases 1 to 86257)  
Martin, S.  
Direct Submission  
Submitted (25-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Feb 25, 2003 this sequence version replaced gi:28208053.  
----- Genome Center

Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:

Em; EMBL; SW; SWISSPROT; Tr; TrEMBL; Mp; WormPEP; Information  
on the WormPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
was generated from part of bacterial clone contigs of human  
Chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping  
Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr1>  
RP11-523M19 is from the library RPCI-11.2 constructed by the group  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBAC3.6.

FEATURES  
source

1.86257  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-523M19"  
/clone\_11b="RPCI-11.2"  
BASE COUNT 29602 a 13951 c 14808 g 27856 t

## ORIGIN

Query Match 91.6%; Score 17.4; DB 9; Length 86257;  
Best Local Similarity 94.7%; Pred. No. 4.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCTTCGACGGGATGCT 19  
|||||  
Db 59984 GGGCTTCGACGGGATGCT 59966

RESULT 16  
AC120698 2/c  
WPCOMMENT

Sequence split into 4 fragments LOCUS AC120698 Accession AC120698

Fragment Name Begin End

AC120698\_0 1 110000  
AC120698\_1 100001 210000  
AC120698\_2 200001 310000  
AC120698\_3 300001 367292

Continuation (3 of 4) of AC120698 Rattus norvegicus clone CH2  
Query Match 91.6%; Score 17.4; DB 2; Length 110000;

Best Local Similarity 94.7%; Pred. No. 4.1e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCTTCGACGGGATGCT 19  
|||||  
Db 102383 GGGCTTCGACGGGATGCT 102365

RESULT 17  
AC117942/c

LOCUS AC117942 142509 bp DNA linear HTG 05-SEP-2002  
DEFINITION Tetradon nigroviridis clone GSTNB-27D21, WORKING DRAFT SEQUENCE, 4  
ordered pieces.

AC117942 GI:22725952  
VERSION HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
KEYWORDS Tetradon nigroviridis  
SOURCE Tetradon nigroviridis  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE  
AUTHORS

1 (bases 1 to 142509)  
Ahter, N., Antonellis, A., Ayale, K., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,  
Brooke, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,  
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,  
Laric, P., Lee-Lim, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,  
Margulies, E.H., Masiello, C., Maekel, B., Mastrian, S.D.,  
McCluskey, J.C., McQuinn, J., Paguirigan, C., Pearson, R.,  
Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schueler, M.G., Sison, C.,  
Santiropp, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,  
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
NSC Comparative Sequencing Initiative

Unpublished  
2 (bases 1 to 142509)  
Green, E.D.

Direct Submission  
Submitted (11-APR-2002) NIH Intramural Sequencing Center, 8717  
Groveomont Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 142509)  
Green, E.D.

Direct Submission  
Submitted (05-SEP-2002) NIH Intramural Sequencing Center, 8717  
Groveomont Circle, Gaithersburg, MD 20877, USA  
On Sep 5, 2002 this sequence version replaced gi:20136917.

## COMMENT

----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: nisc\_zoo@nhgri.nih.gov  
----- Project Information  
Center project name: cvg  
Center clone name: 027D21

The sequence data in this record represents an 'enhanced'  
version of a Phase 2 submission. Specifically, the indicated  
order and orientation of each sequence contig has been  
established using one or more of the following: read-pair  
data from individual subclones, overlaps with neighboring  
clones, alignment with available reference sequence (e.g.,  
human), and/or confirmation by PCR testing. In addition,  
the sequence assembly is based on at least 8x average  
coverage in Q20 bases and has been reviewed to rule out  
gross misassemblies, the low-quality ends of sequence  
contigs have been trimmed away, and each base is associated  
with a Phrap-derived quality score.

----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.950319  
Consensus quality: 141403 bases at least Q40  
Consensus quality: 141846 bases at least Q30

Consensus quality: 142063 bases at least Q20  
 Insert size: 129000; agarose-fp  
 Insert size: 142209; sum-of-ctnigs  
 Quality coverage: 10.93x in Q20 bases; agarose-fp  
 Quality coverage: 9.92x in Q20 bases; sum-of-ctnigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

1 19429: contig of 19429 bp in length  
 \* 19430 19529: gap of unknown length  
 \* 19530 107773: contig of 88244 bp in length  
 \* 107774 107873: gap of unknown length  
 \* 107874 131292: contig of 23419 bp in length  
 \* 131293 131393: gap of unknown length  
 \* 131393 142509: contig of 11117 bp in length.

Location/Qualifiers  
 1. 142509  
 /organism="Tetradodon nigroviridis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:99883"  
 /clone="GSTNB-27D21"  
 /clone\_11b="GSTNB"  
 /note="Genoscope designation: COAB027D21"  
 1. 19429  
 /note="assembly\_fragment  
 clone end:17  
 vector\_side:left"  
 19530. 107773  
 /note="assembly\_fragment"  
 107874. 131292  
 /note="assembly\_fragment"  
 118928. 142509  
 /note="clone overlaps with GenBank Accession Number  
 AC116440 clone GSTNA-49N2 (center project name cvf)"  
 131393. 142509  
 /note="assembly\_fragment  
 clone end:SP6  
 vector\_side:right"

misc\_feature  
 misc\_feature  
 misc\_feature  
 misc\_feature

BASE COUNT 34414 a 35682 c 34612 g 37501 t 300 others  
 ORIGIN

Query Match 91.6%; Score 17.4; DB 2; Length 142509;  
 Best Local Similarity 94.7%; Pred. No. 3.9e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGTCTGACGGGATGCT 19  
 |||||  
 Db 134729 GGGTCTGACGGGATGCT 134711  
 |||||

RESULT 18  
 AC116440 164532 bp DNA linear HTG 23-OCT-2002  
 AC116440/c Tetradodon nigroviridis clone GSTNA-49N2, WORKING DRAFT SEQUENCE, 6  
 LOCUS ordered pieces.  
 AC116440  
 AC116440.3 GI:24270692  
 HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
 KEYWORDS Tetradodon nigroviridis  
 SOURCE Tetradodon nigroviridis  
 ORGANISM Tetradodon nigroviridis  
 Bacteria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Butleostei; Neoteleostei;  
 Acanthomorphi; Acanthopterygii; Percormorpha; Tetradontiformes;  
 Tetradontidae; Tetradontidae; Tetradon.  
 1 (bases 1 to 164532)  
 AUTHORS Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,

# TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Benjamin, B., Blakeley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Cariaga, K., Coleman, B., Dietrich, N.L., Engle, J., Granite, S.,  
 Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L.,  
 Idol, J.R., Karlins, E., Latic, P., Lee-Im, S.-O., Legaapi, R.,  
 Maduro, Q.L., Maduro, V.B., Margulies, E.H., Mastello, C., Maskeri, B.,  
 Mastrian, S.D., McCloskey, J.C., McDowell, J., Paguirigan, C.,  
 Pearson, R., Portnoy, M.E., Prasad, A., Reddi-Dugue, N., Schandler, K.,  
 Schueler, M.G., Sison, C., Stantipod, S., Thomas, J.W., Thomas, P.J.,  
 Touchman, J.W., Vogt, J.L., Walker, M., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 164532)  
 Green, E.D.  
 Direct Submission  
 Submitted (28-MAR-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 3 (bases 1 to 164532)  
 Green, E.D.  
 Direct Submission  
 Submitted (23-OCT-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 On Oct 23, 2002 this sequence version replaced gi:21465411.

----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: http://www.nisc.nih.gov  
 Contact: nisc.zoo@hgrl.nih.gov  
 ----- Project Information  
 Center project name: cvf  
 Center clone name: 049N02

The sequence data in this record represents an 'enhanced'  
 version of a Phase 2 submission. Specifically, the indicated  
 order and orientation of each sequence contig has been  
 established using one or more of the following: read-pair  
 data from individual subclones, overlaps with neighboring  
 clones, alignment with available reference sequence (e.g.,  
 human), and/or confirmation by PCR testing. In addition,  
 the sequence assembly is based on at least 8x average  
 coverage in Q20 bases and has been reviewed to rule out  
 gross misassemblies, the low-quality ends of sequence  
 contigs have been trimmed away, and each base is associated  
 with a Phrap-derived quality score.

----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap, version 0.990319  
 Consensus quality: 163158 bases at least Q40  
 Consensus quality: 163904 bases at least Q30  
 Consensus quality: 163904 bases at least Q20  
 Insert size: 106000; agarose-fp  
 Insert size: 164032; sum-of-ctnigs  
 Quality coverage: 15.47x in Q20 bases; agarose-fp  
 Quality coverage: 10.00x in Q20 bases; sum-of-ctnigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 6 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

1 12551: contig of 12551 bp in length  
 \* 12552 12651: gap of unknown length  
 \* 12652 67722: contig of 55071 bp in length  
 \* 67723 67822: gap of unknown length  
 \* 67823 87024: contig of 19202 bp in length  
 \* 87025 87124: gap of unknown length  
 \* 87125 92799: contig of 5675 bp in length  
 \* 92800 92899: gap of unknown length

```

*          *          *
* 14326 144325: contig of 51426 bp in length
* 144326 144425: gap of unknown length
* 144426 164532: contig of 20107 bp in length.
FEATURES
    source
        1..164532
            /organism="Tetraodon nigroviridis"
            /mol_type="genomic DNA"
            /db_xref="taxon:99883"
            /clone="GSTNA-49N2"
            /clone_lib="GSTNA"
            /note="Genoscope designation: COA049N02"
        misc_feature
            1..24382
                /note="clone overlaps with Genbank Accession Number
                AC117942 clone GSTNB-27D21 (center project name cvg)"
        misc_feature
            1..12551
                /note="assembly_fragment
                clone_end:17
                vector_side:left"
        misc_feature
            12652..67722
                /note="assembly_fragment"
        misc_feature
            67823..87024
                /note="assembly_fragment"
        misc_feature
            87125..92799
                /note="assembly_fragment"
        misc_feature
            92900..144325
                /note="assembly_fragment"
        misc_feature
            144426..164532
                /note="assembly_fragment
                clone_end:596
                vector_side:right"
BASE COUNT      39575 a 41275 c 41473 g 41709 t      500 others
ORIGIN
Query Match      91.6% Score 17.4; DB 2; Length 164532;
Best Local Similarity 94.7%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy      1 GGGTCTGAGCGGATGCT 19
Db      16602 GGGTCTGAGCGGATGCT 16584

RESULT 19
AC105219      177816 bp      DNA      linear      PRI 30-SEP-2002
LOCUS
DEFINITION Homo sapiens chromosome 8, clone RP11-429J17, complete sequence.
ACCESSION      AC105219
VERSION      AC105219.6 GI:23379185
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
    Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
    1 (bases 1 to 177816)
    Birren,B., Nusbaum,C. and Lander,E.
    Homo sapiens chromosome 8, clone RP11-429J17
    Unpublished
    2 (bases 1 to 177816)
    Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
    Anderson,S., Barina,N., Baetien,V., Boguslavsky,L., Boukhalter,B.,
    Brown,A., Camarata,J., Campotien,A., Chang,J., Chazaro,B.,
    Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
    Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S.,
    Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
    Gande,S., Goid,S., Goyette,M., Graham,L., Grand-Pierre,N.,
    Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
    Jones,C., Kamat,A., Karatas,A., Kells,C., Laroque,K.,
    Lamazars,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
    Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
    McCarthy,M., McEwan,P., McKernan,K., McNeesters,R., Meldrim,J.,
    Menueu,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
    Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
    Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

```

```

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (26-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 177816)
    Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
    Barina,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
    Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
    Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
    Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
    Gardyna,S., Goid,S., Graham,L., Grand-Pierre,N., Hagos,B.,
    Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
    Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
    Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
    McCarthy,M., Meldrim,J., Menueu,L., Mihova,T., Mlenga,V.,
    Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
    O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
    Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
    Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
    Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
    Teefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
    Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
    Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 177816)
    Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
    Barina,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
    Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
    Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
    Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
    Gardyna,S., Goid,S., Graham,L., Grand-Pierre,N., Hagos,B.,
    Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
    Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
    Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
    McCarthy,M., Meldrim,J., Menueu,L., Mihova,T., Mlenga,V.,
    Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
    O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
    Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
    Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
    Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
    Teefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
    Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
    Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 30, 2002 this sequence version replaced gi:22759307.
All repeats were identified using RepeatMasker:
    Smit, A.P.A. & Green, P. (1996-1997)
    http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23081
Center clone name: 429_J_17
-----
FEATURES
    source
        1..177816
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="8"

```

```

/map="8"
/clone="RP11-429017"
/clone_lib="RPCT-11 Human Male BAC"
complement(81..161)
repeat_region
/rpt_family="L2"
complement(1950..2022)
repeat_region
/rpt_family="MIR"
complement(2052..2362)
repeat_region
/rpt_family="AluSx"
complement(2382..2491)
repeat_region
/rpt_family="FLAM_C"
complement(3268..3029..3268)
repeat_region
/rpt_family="AluDb"
complement(3992..3929..3992)
repeat_region
/rpt_family="(CTGGGG)n"
complement(4430..4459)
repeat_region
/rpt_family="(CGG)n"
complement(4512..4546)
repeat_region
/rpt_family="(CCC)n"
complement(7830..7860)
repeat_region
/rpt_family="(TCCC)n"
complement(11974..11978)
repeat_region
/rpt_family="GC_rich"
complement(12568..12602)
repeat_region
/rpt_family="GC_rich"
complement(12898..12920)
repeat_region
/rpt_family="GC_rich"
complement(13117..13145)
repeat_region
/rpt_family="GC_rich"
complement(13117..13132)
repeat_region
/rpt_family="GC_rich"
complement(13252..13285)
repeat_region
/rpt_family="GC_rich"
complement(13358..13413)
repeat_region
/rpt_family="GC_rich"
complement(19441..19445)
repeat_region
/rpt_family="GC_rich"
complement(20044..20107)
repeat_region
/rpt_family="(GGA)n"
complement(20089..20093)
repeat_region
/rpt_family="GCCCG)n"
complement(23877..23936)
repeat_region
/rpt_family="GCCCG)n"
complement(27214..27447)
repeat_region
/rpt_family="LIM4"
complement(27448..27548)
repeat_region
/rpt_family="AluDb"
complement(27918..28201)
repeat_region
/rpt_family="AluDb"
complement(28420..28556)
repeat_region
/rpt_family="LIM4"
complement(29100..29192)
repeat_region
/rpt_family="MIR"
complement(29575..29720)
repeat_region
/rpt_family="LIMC5"
complement(29883..30138)
repeat_region
/rpt_family="LIMC4a"
complement(30197..30314)
repeat_region
/rpt_family="FLAM_A"
complement(30523..30676)
repeat_region
/rpt_family="AluSg/x"
complement(30847..30899)
repeat_region
/rpt_family="LIMC4"
complement(30992..31156)
repeat_region
/rpt_family="LIMC4"
complement(31211..31829)
repeat_region
/rpt_family="LIMC4a"
complement(36974..36982)
repeat_region
/rpt_family="LIMC4a"
complement(36992..36998)
repeat_region
/rpt_family="C-rich"
complement(39443..39536)

```

```

repeat_region 41500..41662
/rpt_family="LTR55"
repeat_region 41771..41867
/rpt_family="LTR55"
repeat_region 41921..42060
/rpt_family="L2"
repeat_region 42087..42176
/rpt_family="L2"
repeat_region 42633..42770
/rpt_family="AluDb"
repeat_region 42771..43069
/rpt_family="AluSg"
repeat_region 43070..43239
/rpt_family="AluDb"
repeat_region 43240..43543
/rpt_family="AluSx"
repeat_region 43718..43900
/rpt_family="LIM3A"
repeat_region 43903..44255
/rpt_family="MLT11"
repeat_region complement(44256..44634)
/rpt_family="MSTA"
repeat_region 44635..44661
/rpt_family="MLT11"
repeat_region 44766..45629

Query Match 91.6% Score 17.4; DB 9; Length 177816;
Beet Local Similarity 94.7%; Pred. No. 3.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCTGCGAGCGGATGCT 19
Db 128432 GGGCTGCGAGCGGATGCT 128450

RESULT 20
AC104885/C
LOCUS
DEFINITION Mus musculus clone RP23-125P19, WORKING DRAFT SEQUENCE, 6 ordered
pieces
AC104885
AC104885.3 GI:24182334
VERSION
KEYWORDS HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
SOURCE
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 200315)
Birren,B., Nusbäum,C. and Lander,E.
Mus musculus, clone RP23-125P19
Unpublished
2 (bases 1 to 200315)
Birren,B., Linton,L., Nusbäum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Baetien,V., Bogunavskiy,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatae,A., Kells,C., Larkov,K.,
Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
Meneu,L., Minova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Neubou,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Roman,J., Roerich,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamae,J., Tefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

COMMENT

Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 200315)  
Britten, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Camata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gori, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,  
Mathews, C., McCarthy, M., Meldrum, J., Meneses, J., Mihova, T.,  
Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, D.,  
Peterson, K., Phunhkhang, P., Pierre, N., Raymond, C., Retta, R.,  
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R.,  
Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N.,  
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (21-OCT-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Oct 21, 2002 this sequence version replaced gi:19881860.  
All repeats were identified using RepeatMasker:  
Smt, A. F. A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WtBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L18985  
Center clone name: 125\_P\_19  
----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 199225 bases at least Q40  
Consensus quality: 199225 bases at least Q30  
Consensus quality: 199621 bases at least Q20  
Insert size: 19300; agarose-fp  
Insert size: 199815; sum-of-coverage  
Quality coverage: 10.4 in Q20 bases; sum-of-coverage  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1  
36119: contig of 36019 bp in length  
36019: contig of 100 bp  
36020: contig of 15755 bp in length  
51875: gap of 100 bp  
51875: contig of 23956 bp in length  
51975: gap of 100 bp  
75931: contig of 28887 bp in length  
76031: gap of 100 bp  
104917: contig of 73293 bp in length  
104918: gap of 100 bp  
178310: contig of 73293 bp in length  
178311: gap of 100 bp  
178411: contig of 21905 bp in length.  
Location/Qualifiers  
1. 200315  
/organism="Mus musculus"

misc\_feature  
1..36019  
/note="assembly\_fragment  
vector side:left"  
clone\_end:SP6  
36120..51874  
/note="assembly\_fragment"  
51975..75930  
/note="assembly\_fragment"  
76031..104917  
/note="assembly\_fragment"  
105018..178310  
/note="assembly\_fragment"  
178411..200315  
/note="assembly\_fragment  
vector side:right"  
clone\_end:T7

BASE COUNT 53188 a 46038 c 47231 g 53356 t 502 others

ORIGIN

Query Match 91.6%; Score 17.4; DB 2; Length 200315;  
Best Local Similarity 94.7%; Pred. No. 3.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OR 1 GGGCTGACGGCGATGCT 19  
Db 154350 GGGCTGACGGCGATGCT 154332

RESULT 21  
AC11254  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-32017, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 6 unordered pieces.  
AC11254  
VERSION  
AC11254.4 GI:30578635  
KEYWORDS  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
SOURCE  
Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 246937)  
Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D.,  
Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, B.,  
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Centes, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Devila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denon, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C. M., Gabist, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,  
Gunnarsson, P., Haland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,  
Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpthy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorenshewa, L., Louised, H., Lozada, R. J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,

Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
 Mlosovlevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
 Narkervic, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwokoileme, O., Okwuonu, G., Olarnunsaigon, A., Pal, S., Parks, K.,  
 Paternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,  
 Plapp, F., Poldexter, A., Popovic, D., Primus, B., Pu, L.,  
 Puzo, M., Quirz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R.,  
 Rell, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rivers, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S.,  
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
 Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,  
 Williams, G., Willson, R., Wleczek, R., Wood, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstein, G. and Gibbs, R.A.

Unpublished  
 2 (bases 1 to 246937)  
 Direct Submission  
 Worley, K.C.  
 Direct Submission  
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 246937)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On May 13, 2003 this sequence version replaced gi:23269171.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: CH230-32017  
 Center clone name: GLMX  
 ----- Summary Statistics  
 Assembly program: Atlas 3.0/  
 Consensus quality: 232316 bases at least Q40  
 Consensus quality: 236042 bases at least Q30  
 Consensus quality: 238624 bases at least Q20  
 Estimated insert size: 248705; sum-of-contigs estimation  
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----  
 \* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 6 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will  
 \* be preserved.

1 235507: contig of 235507 bp in length  
 \* 235508: gap of unknown length  
 \* 235608: contig of 1065 bp in length  
 \* 236773: gap of unknown length  
 \* 236773: contig of 1047 bp in length  
 \* 237820: gap of unknown length  
 \* 237920: contig of 2615 bp in length  
 \* 240535: gap of unknown length  
 \* 240635: contig of 2126 bp in length  
 \* 242761: gap of unknown length  
 \* 242861: gap of unknown length  
 \* 242861: contig of 4077 bp in length.

FEATURES  
 source  
 1..246937  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-32017"

misc\_feature  
 1..1422  
 /note="wgs contig"

misc\_feature  
 232662..235507  
 /note="wgs contig"

BASE COUNT 67730 a 56226 c 53928 g 62488 t 6565 others

ORIGIN

Query Match 91.6%; Score 17.4; DB 2; Length 246937;  
 Best Local Similarity 94.7%; Pred. No. 3.7e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTCTGCACGGGATGCT 19  
 |||||  
 DB 105519 GGGTCTGCACGGGATGCT 105537

RESULT 22  
 LOCUS AX008662 19 bp DNA linear PAT 06-SEP-2000  
 DEFINITION Sequence 15 from Patent WO9966037.  
 ACCESSION AX008662  
 VERSION AX008662.1 GI:9996186  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 synthetic construct  
 synthetic construct  
 artificial sequences.

REFERENCE  
 1 Renzi, P.  
 Antisense oligonucleotides for treating or preventing atopic  
 diseases and neoplastic cell proliferation  
 Patent: WO 9966037-A 15 23-DEC-1999;  
 JOURNAL RENZI PAOLO (CA); RECH EXPERTISES ET DEV MEDICAU (CA)  
 FEATURES  
 source  
 1..19  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 /note="Antisense oligonucleotide inhibiting the common  
 subunit of IL-3, IL-5 and GM-CSF human receptor"

BASE COUNT 2 a 3 c 8 g 6 t

ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GTCTGCACGGGATGCT 19  
 |||||  
 DB 1 GTCTGCACGGGATGCT 17

RESULT 23  
 AX671174 19 bp DNA linear PAT 27-MAR-2003  
 LOCUS AX671174



DEFINITION Sequence 14 from Patent WO03004511.  
 ACCESSION AX671182  
 VERSION AX671174.1 GI:29329630  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 REFERENCE 1  
 AUTHORS Renzi, P., Allam, M. and Allakhverdi, Z.  
 TITLE Methods for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
 JOURNAL Patent: WO 03004511-A 14 16-JAN-2003;  
 Topigen Pharmaceuticals Inc (CA)  
 FEATURES  
 source 1..19  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 /note="Sequence is completely synthesized"  
 BASE COUNT 2 a 3 c 8 g  
 ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 GGTCTGCAGCGGATGCT 19  
 Db 1 GGTCTGCAGCGGATGCT 17

RESULT 24  
 LOCUS AX671181 19 bp DNA linear PAT 27-MAR-2003  
 DEFINITION Sequence 21 from Patent WO03004511.  
 ACCESSION AX671181  
 VERSION AX671181.1 GI:29329637  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 REFERENCE 1  
 AUTHORS Renzi, P., Allam, M. and Allakhverdi, Z.  
 TITLE Methods for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
 JOURNAL Patent: WO 03004511-A 21 16-JAN-2003;  
 Topigen Pharmaceuticals Inc (CA)  
 FEATURES  
 source 1..19  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 /note="Sequence is completely synthesized"  
 misc\_feature 9  
 (name="N" corresponds to 2,6-diaminopurine nucleoside (DAP) "  
 misc\_feature 15  
 (note="N" corresponds to 2,6-diaminopurine nucleoside (DAP) "  
 BASE COUNT 0 a 3 c 10 g 4 t 2 others  
 ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 19;  
 Best Local Similarity 89.5%; Pred. No. 1.9e+03;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGTCTGCAGCGGATGCT 19  
 Db 1 GGTCTGCAGCGGATGCT 19

RESULT 25  
 LOCUS AX671182

LOCUS AX671182 19 bp DNA linear PAT 27-MAR-2003  
 DEFINITION Sequence 22 from Patent WO03004511.  
 ACCESSION AX671182  
 VERSION AX671182.1 GI:29329638  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 REFERENCE 1  
 AUTHORS Renzi, P., Allam, M. and Allakhverdi, Z.  
 TITLE Methods for increasing in vivo efficacy of oligonucleotides and inhibiting inflammation in mammals  
 JOURNAL Patent: WO 03004511-A 22 16-JAN-2003;  
 Topigen Pharmaceuticals Inc (CA)  
 FEATURES  
 source 1..19  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 /note="Sequence is completely synthesized"  
 modified\_base 9 /mod\_base=1  
 modified\_base 15 /mod\_base=1  
 BASE COUNT 0 a 3 c 10 g 4 t 2 others  
 ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 19;  
 Best Local Similarity 89.5%; Pred. No. 1.9e+03;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGTCTGCAGCGGATGCT 19  
 Db 1 GGTCTGCAGCGGATGCT 19

RESULT 26  
 LOCUS AE004803 10225 bp DNA linear BCT 19-FEB-2003  
 DEFINITION Pseudomonas aeruginosa PA01, section 364 of 529 of the complete genome.  
 ACCESSION AE004803 AE004091  
 VERSION AE004803.1 GI:9950035  
 KEYWORDS  
 SOURCE Pseudomonas aeruginosa PA01  
 ORGANISM Pseudomonas aeruginosa PA01  
 REFERENCE 1 (bases 1 to 10225)  
 AUTHORS Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kae, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Sater, M.H., Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
 JOURNAL Nature 406 (6799), 959-964 (2000)  
 MEDLINE 10984043  
 PUBMED 20437337  
 REFERENCE 2 (bases 1 to 10225)  
 AUTHORS Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kae, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Sater, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA  
 REFERENCE 3 (bases 1 to 10225)

AUTHORS  
CONSRM  
TITLE  
JOURNAL

Pseudomonas aeruginosa Community Annotation Project (PseudocAP)  
Direct Submission  
Submitted (04-FEB-2003) Department of Molecular Biology and  
Biochemistry, Simon Fraser University, 8088 University Dr.,  
Burnaby, British Columbia V5A 1S6, Canada

## COMMENT

This represents the February 3, 2003 version of the continually  
updated, reviewed, Pseudomonas aeruginosa PAO1 genome annotation,  
from PseudocAP (see <http://www.pseudomonas.com> for latest updates  
and links to alternate annotations). PseudocAP is coordinated by  
Flora S.L. Brinkman (Simon Fraser University, Canada) and Robert  
E.W. Hancock (University of British Columbia, Canada). We welcome  
submission through [www.pseudomonas.com](http://www.pseudomonas.com) of any proposed changes.

'Protein name confidence' is used to rate our confidence of the  
accuracy of the protein name.  
Class 1: Function experimentally demonstrated in P. aeruginosa.  
Class 2: Function of highly similar gene experimentally  
demonstrated in another organism (and gene context consistent  
in terms of pathways its involved in, if known).  
Class 3: Function proposed based on presence of conserved amino  
acid motif, structural feature or limited sequence similarity  
to an experimentally studied gene.  
Class 4: Homologs of previously reported genes of unknown function,  
or no similarity to any previously reported sequences.

## FEATURES

## source

Location/Qualifiers  
1..10225  
/organism="Pseudomonas aeruginosa PAO1"  
/mol\_type="genomic DNA"  
/strain="PAO1"  
/db\_xref="taxon:208964"  
complement(87..1112)  
/locus\_tag="PA3858"  
/note="synonym: aapJ"  
complement(87..1112)  
/locus\_tag="PA3858"  
/note="Protein name confidence: Class 3 (function proposed  
based on presence of conserved amino acid motif,  
structural feature or limited sequence similarity to an  
experimentally studied gene); Subcellular localization:  
periplasmic; Subcellular localization confidence: Class 2"  
/codon\_start=1  
/transl\_table=11  
/product="probable amino acid-binding protein"  
/protein\_id="AAG07245.1"  
/db\_xref="GI:9950036"  
/translation="MTIMKGTWLLLAANVCLASATPAEAGATLDAVKRGVYCGVSDG  
LGFSYADAKGRFGKLDVDCALAAALFGDAEKVFPPLSTERSALOSGTNLLS  
BNTSWTSRRDTLGIHFGVYFDCQGLVNRKLGVSARBELKDATVIGOSGTNELL  
VADYFRAGIRIKAVLIDBSYEKSAALTEGQCDVSSDQSLYARLKLAKREQVYL  
PEVISEKPEPAVRQGDDEAFNIVRWTLGLNASELGYVTSNVERQADSNDPVAR  
LIGSGDAGKDLQLPEDWVQWVRQVNGVEIFARNVGDGSLKMRGLNAGNLGL  
HYAPPIR"  
1409..2056  
/locus\_tag="PA3859"  
1409..2056  
/locus\_tag="PA3859"  
/note="Protein name confidence: Class 3 (function proposed  
based on presence of conserved amino acid motif,  
structural feature or limited sequence similarity to an  
experimentally studied gene)"  
/codon\_start=1  
/transl\_table=11  
/product="probable carboxylesterase"  
/protein\_id="AAG07246.1"  
/db\_xref="GI:9950037"  
/translation="WSEPLILDAPNADACIITLHIGAGNRDTPKPAVALQVLT  
FTLPQASQAVTNGWMPWSTYDIASPARAIDEDQANASADQVLAIDQRAKGI  
AAERITLAFSGAGAVLHARRRYAQPGVLAISTAVPTDQALDELRHRRITVLH  
LHGSDVDVPLGAAHDALQAGVEVGMHDYPMGHEVSLIEIHIDIGAMKRL"  
complement(2072..3970)

## CDS

/locus\_tag="PA3860"  
complement(2072..3970)  
/locus\_tag="PA3860"  
/note="Protein name confidence: Class 3 (function proposed  
based on presence of conserved amino acid motif,  
structural feature or limited sequence similarity to an  
experimentally studied gene); Subcellular localization:  
cytoplasmic; Subcellular localization confidence: Class 2"  
/codon\_start=1  
/transl\_table=11  
/product="probable AMP-binding enzyme"  
/protein\_id="AAG07247.1"  
/db\_xref="GI:9950038"  
/translation="WNPSPIAGSDIEALERVLPQRDPSSTYELLQSRHQR  
IALSCLHGSAAEPRIRISYAEIPAVYOTANALHRLGDSHQAISFLPQLPQHY  
IWGEAGIIVAINPILPEPIHAEILASNTAVLTLPAPFGTDLMQKAGLAPLPE  
LYAIWVDPANLLPAPORBALKAQRGLPESVLPDILLDCPADRLSEGRADHPDV  
ASYHTGGTCTTPKLAHSHENRPAAMEIKMLANDYGVDDVLLGLPLFYNGVMWG  
LAPHRGAQVILAPQGRNPTLLIQDFMKLVERIVTSFSGVPTTYALALVQPSDGRD  
LSLRPFLCGAAPMPVELIRQFEARTGLKVEGLTEGTCGSCNRRGERRGSGIG  
LRLPYCQVAVVLDGEGYLRDAPNVEGNCLGPTVFKGYLQODNRDRIWDGMP  
MTGDLRIDEDGYIWLGRSKDLIRGGHNDPQMIIEALHRAVLAALAAVGRDAK  
AGELPVAYIOLKPGASASEELAEHARHVRPAVAKDILISMPTVYAGKFKPA  
LRLAIVRLEESRRIAEIDIRVAVADERHQLAHLHVPALDRQAALELLGVA  
LNYRLHAV"

## gene

## CDS

/locus\_tag="PA3861"  
4584..5777  
/gene="rhl"  
/locus\_tag="PA3861"  
/note="Protein name confidence: Class 2 (high similarity  
to functionally studied protein)"  
/codon\_start=1  
/transl\_table=11  
/product="ATP-dependent RNA helicase RhlB"  
/protein\_id="AAG07248.1"  
/db\_xref="GI:9950039"  
/translation="MVEPQEGKTRFHDNLAPSLMHAHIDGFPYCTPIQOVLGFTL  
RGDAIGAGTGTGATLAFISITLQLOTPPKERVYGEPRALIAPTRELVAQIAK  
DAALTYGTGNTWTFVCGMPDFOKOLBARFDIIVATGRLLDFNRQEVHLDVY  
EVMVDEADRLMDGFTIPQVQITIRQPHKERTLLPSAFTPDVWVLAQWVDA  
IVELIPEENVASDIVEQHYAVASDKLNLNVAQNMEVMPFARKDEVRRIEER  
LTKDISAQMGSQVDPQHKIRITLGEFEGKIRLVAVADVAGRHGIIHDGISHVINFPL  
PEDPDVYHRIGRTRAGASGTSISFAGEDDAPALPIEBLGRKTCMEMPABLKP  
VPRKH"  
5938..6885  
/locus\_tag="PA3862"  
5938..6885  
/locus\_tag="PA3862"  
/note="Protein name confidence: Class 4 (homologs of  
previously reported genes of unknown function, or no  
similarity to any previously reported sequences)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAG07249.1"  
/db\_xref="GI:9950040"  
/translation="WSAATPLIVQAAEAOILARIIDVLOAMROPLFLDAAGALOPAO  
QLVEFPARGSPFINTLGLVLADEOYGVTSYTRISQEPPLNTATMLSMQTPILL  
CDARLTARTARNTAAVAVDALPAEACRLALISGPAHAAHLVTKLRMQRVRA  
SPCIDRELQSLRAIDIPRAAGSLEELADADVILCTSSARAVIDRQURPALVT  
SISTNAPRAHVEPSLAAMDVYCDYRTTGDSGEMILIAEHOIGMSPEAIRGLAEL  
LSAQAQPREYRRPAFRFISIGLBDVALNANLYLRQNG"  
6908..8035  
/locus\_tag="PA3863"  
6908..8035  
/locus\_tag="PA3863"  
/note="Protein name confidence: Class 4 (homologs of  
previously reported genes of unknown function, or no  
similarity to any previously reported sequences)"  
/codon\_start=1  
/transl\_table=11

## gene

/product="hypothetical protein"  
/protein\_id="AA07250.1"  
/db\_xref="GI:9950041"  
/translation="MEADYLVICGAGAGNGVLSAHGRVVLREAOQCHYSTGR  
SAHTTVAYQKRLDAEQASIVPLARDKVGATYDADIDTDLAQYLRIRQ  
YESGVALVQKRLDAEQASIVPLARDKVGATYDADIDTDLAQYLRIRQ  
NOGQVLCHEALVLRVGVAMEVRCDSAYAAVYVNAAGMCDIAGIACVPLGLQ  
PKRRSAFIFAPPIDGIDHMPVLVLDSESYLAKPKPAGMLGSPANADVEAHADYQPEO  
LDIATGMVLEATLTITRREPHTAGRSFVADSDIYAGYAANAESEFWAAACGCG  
IOTSAAMEASALIRHOPRAHLREHLDENMLSPRLSP"  
8098..8730  
/locus\_tag="PA3864"  
8098..8730  
/locus\_tag="PA3864"  
/note="Protein name confidence: Class 4 (homologs of  
previously reported genes of unknown function, or no  
similarity to any previously reported sequences)"  
/codon\_start=1  
/transl\_table=11

Query Match 89.5%; Score 17; DB 1; Length 10225;  
Best Local Similarity 100.0%; Pred. No. 8.6e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCTGCAGCGGATG 18  
|||||  
Db 8970 GGCTGCAGCGGATG 8986

RESULT 27  
AX066133 436 bp DNA linear PAT 24-JAN-2001  
LOCUS Definition Sequence 37 from Patent WO0100804.  
ACCESSION AX066133  
VERSION AX066133.1 GI:12543845  
KEYWORDS  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinomycetales;  
Corynebacteriaceae; Corynebacterium.  
REFERENCE 1  
Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O., Habermann, G.,  
Lee, H. S. and Kim, H. J.  
corynebacterium glutamicum genes encoding stress, resistance and  
tolerance proteins  
Patent: WO 0100804-A 37 04-JAN-2001;  
JOURNAL BASF AKTIENGESELLSCHAFT (DE)  
FEATURES  
source  
1..436  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1718"  
101..2436  
/note="FRX02282"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CA26305.1"  
/db\_xref="GI:12543846"  
/translation="MQESSRDNFQVDLGVLDLSRHYSPRYVVELLQNAVDACT  
ARSGEGYEPSPRIRPVTKDRATFSLVDNGTGLTVAQEARRELLATVGRTSKDEFGI  
QREGRLGQFG"  
BASE COUNT 81 a 102 c 138 g 115 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 6; Length 436;  
Best Local Similarity 94.4%; Pred. No. 2.5e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCTGCAGCGGATG 19  
|||||  
Db 401 GGCTGCAGCGGATG 418

RESULT 28  
AB02081540/c 778 bp DNA linear PRI 14-APR-2000  
LOCUS Homo sapiens gene for DMBT1, exon 40, complete cds.  
ACCESSION AB020815  
VERSION AB020815.1 GI:4996276  
KEYWORDS DMBT1.  
SEGMENT 40 of 40  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
Takeshita, H., Sato, M., Shiwa, H. O., Sema, S., Sakurada, A.,  
Hoshi, M., Hayashi, Y., Tagawa, Y., Ayabe, H. and Horii, A. A.,  
Expression of the DMBT1 gene is frequently suppressed in human lung  
cancer  
Jpn. J. Cancer Res. 90 (9), 903-908 (1999)  
JOURNAL Upr. J. Cancer Res. 90 (9), 903-908 (1999)  
MEDLINE 20017478  
PUBMED 10551316  
2 (bases 1 to 778)  
AUTHORS Horii, A.  
TITLE Direct Submission  
JOURNAL Submitted (06-DEC-1998) Akira Horii, Tohoku University School of  
Medicine, Department of Molecular Pathology, 2-1 Seiryomachi,  
Aoba-ku, Sendai, Miyagi 980-8575, Japan  
(E-mail: horii@mail.cc.tohoku.ac.jp, Tel: 81-22-717-8042,  
Fax: 81-22-717-8047)  
FEATURES  
source  
1..778  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
join(AB020812.1:179..239,AB020813.1:78..107,  
AB020814.1:82..129,AB020815.1:102..149,AB020816.1:74..121,  
AB020817.1:89..136,AB020818.1:74..397,AB020819.1:84..113,  
AB020820.1:79..120,AB020821.1:90..413,AB020822.1:84..109,  
AB020823.1:85..117,AB020824.1:85..406,AB020825.1:84..113,  
AB020826.1:85..117,AB020827.1:85..406,AB020828.1:81..110,  
AB020829.1:78..110,AB020830.1:84..407,AB020831.1:83..112,  
AB020832.1:81..113,AB020833.1:79..402,AB020834.1:84..113,  
AB020835.1:73..105,AB020836.1:81..408,AB020837.1:84..107,  
AB020838.1:78..116,AB020839.1:92..415,AB020840.1:77..103,  
AB020841.1:82..114,AB020842.1:79..232,AB020843.1:88..290,  
AB020844.1:83..403,AB020845.1:88..108,AB020846.1:84..113,  
AB020847.1:80..230,AB020848.1:81..289,AB020849.1:84..268,  
AB020850.1:86..495,74..350)  
/gene="DMBT1"  
join(AB020812.1:179..239,AB020813.1:78..107,  
AB020814.1:82..129,AB020815.1:102..149,AB020816.1:74..121,  
AB020817.1:89..136,AB020818.1:74..397,AB020819.1:84..113,  
AB020820.1:79..120,AB020821.1:90..413,AB020822.1:84..109,  
AB020823.1:85..117,AB020824.1:85..406,AB020825.1:84..110,  
AB020826.1:85..115,AB020827.1:83..406,AB020828.1:81..110,  
AB020829.1:78..110,AB020830.1:84..407,AB020831.1:83..112,  
AB020832.1:81..113,AB020833.1:79..402,AB020834.1:84..113,  
AB020835.1:73..105,AB020836.1:81..408,AB020837.1:84..107,  
AB020838.1:78..116,AB020839.1:92..415,AB020840.1:77..103,  
AB020841.1:82..114,AB020842.1:79..232,AB020843.1:88..290,  
AB020844.1:83..403,AB020845.1:88..108,AB020846.1:84..113,  
AB020847.1:80..230,AB020848.1:81..289,AB020849.1:84..268,  
AB020850.1:86..495,74..350)  
/gene="DMBT1"  
/codon\_start=1  
/product="DMBT1"  
/protein\_id="BA078577.1"  
/db\_xref="GI:4996276"  
/translation="MGISTVLEMCILMGQVLSSTGKIPRTTYASLIPEBVPDITV  
AESPPSELTSTVAEGSPISLESTVABGSLIPESLTLESTVAGSGSLAL  
RLVNGDRCQGRVYLIRGSGWTVCDSDWTQDANVAVCRQAGMAAGNAPFGG  
SGPIALDDVRCGHESTYLMSCPHNGVLSHNGGHEAGVCSAQAQPSLTPSPSPWR  
ISPPVPEGSESSIALRLVNGDRCQGRVYLIRGSGWTVCDSDWTQDANVAVCRQAG  
CGWMSAPGNAQFGGSGPIALDDVRCGHESTYLMSCPHNGVLSHNGGHEAGVCSAQAQPSLTPSPSPWR

```

-35_signal
-10_signal
RBS
gene
CDS

/organism="Paedomonas aeruginosa"
/mol_type="genomic DNA"
/strain="PAO1"
/db_xref="taxon:287"
94. .99
117. .122
148. .151
160. .1437
/gene="oprQ"
160. .1437
/gene="oprQ"
/codon_start=1
/translation_table=11
/product="OprE3"
/protein_id="BAA22267.1"
/db_xref="GI:2425172"
/translation="MLKKRICLALGATATASQAMANDQEAKEFVEDSHLDLFFRCY
ISDYGKRGSDAKEMGQAAATATPTSGTQQTIVGVDAFGLVAVRLDGGKRSQAADI
DPKQSGSGAADDLKSQAAVAFRIISNTVLKYGDOMPSLPVLSYDNRSLIPESGCT
LINSKEGLEELNAGRPTAESRSKSAEGRDGGLKSIIVFGAKYAFTHFNASLYASDV
EDLTKYQYINLANTTIPLOADQSLNPDENGVRGRLTSPDAONNGNRDNHLSLAAYV
TIDGHTFMINQRNTGDTGNYKGYONAGSIGCGCTTIWLANSYNSPFNAEDRSQNV
SYALDPAKGVPGVLTAVAVRGDNTIKATSTSKEREITFNOVOYVQVQSGPAKDLTLR
LRSSFLRVNSDARSYNDNGNEIRAPAFVEPSPV"
223. .1434
/gene="oprQ"
/product="OprE3"
/evidence=experimental
BASE COUNT      332 a      485 c      421 g      230 t
ORIGIN

Query Match      86.3%; Score 16.4; DB 1; Length 1468;
Best Local Similarity 94.4%; Pred. No. 2,2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

oy	2	GGCTGCAGCGGGATGTT	19
Db	870	GGCCTGCAGCGGGATGTT	853

RESULT 30  
AK120193  
\*\*\*\*\*  
1994 JUN DWN 140000  
PMT 11 MAY 2001

DEFINITION Sequence 109 from Patent EPI108790.  
 ACCESSION AX120193  
 VERSION AX120193.1 GI:14036908

**SOURCE ORGANISM**  
Corynebacterium glutamicum  
Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Actinomycetaceae; Actinomyces

**REFERENCE  
AUTHORS**

1  
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.

**FEATURES**  
**JOURNAL**  
 Patent: EP 1108790-A 109 20-JUN-2001;  
 KYOWA HAKKO KOGYO CO., LTD. (JP)  
 Location/Qualifiers  
 1 1004

```

375 /organism="Corynebacterium glutamicum"
376 /mol_type="genomic DNA"
377 /db_xref="taxon:1718"
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
```

Query Match	Score	DB 6	Length
86.3%	16.4	6	1824
84.8%	16.2	6	1840

Matches	17;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	2	GGCTGCAGCGGATGCT	19						

Db 301 GGTCTGCAGCGGNAAGT 318



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLINK at: <http://image.lnl.gov>  
Series: ITAL Plate: 33 Row: 1 Column: 22  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

## FEATURES

## Source

## CDS

Location/Qualifiers  
1..2097  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:23427 IMAGE:4654320"  
/tissue\_type="Skin, melanoma, amelanotic"  
/clone\_lib="NIH MGC\_41"  
/lab\_host="DH10B-R"  
/note="vector: pOTB7"  
151..1914  
/codon\_start=1  
/product="Similar to RIKEN CDNA 061020102 gene"  
/protein\_id="AAH15649.1"  
/db\_xref="GI:15990515"  
/translation="MSQMLHIFIPKNTVAGCLNEORLLGXCDVSVIWKQAKRA  
RAVLASSLIFRDLFSNGSKAFELPSVPPACFOQLISFCTGRLTNTASQVMT  
TAGFLQIQLIVERGTDLMFKVSSPRCDQTAVIDAGSPQSPCNQLOPAAAAAPV  
VSPSVPLILTRVKHEAMELPAGGLAPKRPLEPRDGVAVAAAGAAVAGTAPLK  
PRSVYGVPLATLPGIQOMPYPQGERTSRQASLPTTDSPTSYNEDEDEDDAD  
TMVEROYGMVTKASGSYAVOEKRPVLESSTCYLIRDDVALPASLISOYRCHP  
KLISRGDGEKLELVAGSGVYTRQLNCHCAGVKKVLLRLALPPRNTLANS  
CGTGRSTSDSRKPLDSRYLVANVAKLQONFAPFSKEMVTHADCTNRRYRKA  
WLPIKSMLEGEVETRTVMSAAASVPLDEFPFAAQVEQRIYARRDQDAITVA  
LRTDANVNDLSAANPAPDAGEVAGSVIOEVAAPEPLPADGSPPOPEGGGGP  
SRPOTPAARPEGTAGTL"

BASE COUNT 394 a 740 c 658 g 305 t

Query Match 86.3%; Score 16.4; DB 9; Length 2097;  
Best Local Similarity 94.4%; Pred. No. 2.1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OR 1 GGCTTCGACGGCGGATCG 18  
|||||  
Db 664 GGCTCAGCAGCGGATCG 647

## RESULT 34

LOCUS MLCPMOA 4690 bp DNA linear BCT 15-JUN-2001  
DEFINITION Methylococcus capsulatus particulate methane monooxygenase subunit  
(pmoC1), particulate methane monooxygenase 27 kDa subunit (pmoA),  
and particulate methane monooxygenase 45 kDa subunit precursor  
(pmoB) genes, complete cds.

ACCESSION L40804  
VERSION L40804.2  
KEYWORDS GI:14456718

SOURCE Methylococcus capsulatus  
Methylococcus capsulatus  
Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;  
Methylococcaceae; Methylococcus.

## REFERENCE

AUTHORS 1 (bases 1 to 4690)  
Semrau, J.D., Chistoserdov, A., Lebron, J., Costello, A., Davagnino, J.,  
Kenna, B., Holmes, A.J., Finch, R., Murrell, J.C. and Lidstrom, M.E.  
Particulate methane monooxygenase genes in methanotrophs

J. Bacteriol. 177 (11), 3071-3079 (1995)  
95286486

## MEDLINE

## PUBMED

2 (bases 1432 to 4689)  
Lidstrom, M.E.  
Submitted (14-MAR-1997) Chemical Engineering, University of  
Washington, Seattle, WA 98195, USA

## JOURNAL

REFERENCE 3 (bases 1 to 4690)  
Stolyar, S.E. and Lidstrom, M.E.  
Direct Submission

JOURNAL Submitted (06-JUN-2000) Chemical Engineering, University of  
Washington, Seattle, WA 98195, USA  
REMARK Sequence updated by submitter  
COMMENT On Jun 15, 2001 this sequence version replaced gi:1894817.  
FEATURES

## Source

## gene

## CDS

1..4690  
/organism="Methylococcus capsulatus"  
/mol\_type="genomic DNA"  
/strain="Bath"  
/db\_xref="taxon:414"  
1557..2339  
/gene="pmoC1"  
1557..2339  
/gene="pmoC1"  
/note="pmoC1"  
/note="pmoC1; putative"  
/codon\_start=1  
/transl\_table=1  
/product="particulate methane monooxygenase subunit"  
/protein\_id="AAB49820.1"  
/db\_xref="GI:1894818"  
/translation="MAATTTGGAAGAABAPLLDKKLTFLALAYTVFYLMVRYEGVYG  
WSAGLDSRAPRPFETVMNPLYTEIVLIVTASILMGYLMKTRDRNLALTPRELRN  
FTHLVWLYAYAKLYKASVYTBODGTTHQTIYDPTPETHIEFLYLYIYITGP  
AAFLYAKTRLEPFAKGISLPTLVVGFPMILPNVGLNMGHTFMSEELFVALHYG  
FVIGWALAVMGTLTQTFYRPAQGGGLCEAVDGLAK"  
2482..3237  
/gene="pmoA"  
2482..2486  
/gene="pmoA"  
/note="putative"  
2494..3237  
/gene="pmoA"  
/note="pmoA"  
/codon\_start=1  
/transl\_table=1  
/product="particulate methane monooxygenase 27 kDa  
subunit"  
/protein\_id="AAB49821.1"  
/db\_xref="GI:790831"  
/translation="MSAAQSAVRSHAEVQVSRITDMALFVPEVIVSYHIAHLT  
MGMDPMSDKDRRLNTYTPYIVLTPANQSLMTRYLPMGATCVGLLGEWT  
NRYTFMGWTFPLNFVPPASLVPGAILDVLMLSSSYFLTAIVGAKGLIYRGN  
WPIIAPLHVPEVNNGLMSIADIQGYVYRTGPEYIRWEKGLRTFGKQAVASAP  
FSAFMSILYFMHPIGRWFSNERFLOST"  
3334..4588  
/gene="pmoB"  
3334..3337  
/gene="pmoB"  
/note="putative"  
3344..4588  
/gene="pmoB"  
/note="pmoB"  
/codon\_start=1  
/transl\_table=1  
/product="particulate methane monooxygenase 45 kDa subunit  
precursor"  
/protein\_id="AAB49822.1"  
/db\_xref="GI:790832"  
/translation="MKTIKRIAKMSAIGLSNAVATFYPASASAGKSGQAAPFMN  
RTIHVYDLSWSKKEVKNIEVEIKGFHVEGPEYVDEPDVAFNLGMPKVFIRK  
SYIGQIVPRVRLIEKTYDFRVVLARRGDWHVITMNVAGGSPITGKVIYVE  
GSMSEFNPVTTLTGQTVLENNEGNTYFWHAFWPAIVAMIGWSRRPIFIRLLM  
VDAGRAELVATDRKVMGFLAATILIVNAMSASAKYITIPLAGTWRGKPKLE  
LPARTVVAEGSDATYRVGRAMKMLITTHNGNSRILGEFTASVFLDSVDYKDDT  
GYPEDLAEGSLSYSDNSPLAPGRTETVVDYVASDAANAEVRLSDIITDPSRFAGLLE  
FFDATGNRQVVQIDAPLPSFM"

## gene

## RBS

## CDS

sig\_peptide  
mat\_peptide

BASE COUNT 900 a 1400 c 1438 g 952 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 1; Length 4690;  
Best Local Similarity 94.4%; Pred. No. 1.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCTGCGAGCGGATGCT 19  
Db 4141 GGCTGCGAGCGGATGCT 4124

RESULT 35  
A90822/c  
LOCUS A90822 5802 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 1 from Patent WO9830687.  
ACCESSION A90822  
VERSION A90822.1 GI:6739232  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 5802)  
AUTHORS Poustka, A. and Mollenhauer, J.  
TITLE PROTEIN CONTAINING AN SRCR DOMAIN  
JOURNAL Patent: WO 9830687-A 1 16-JUL-1998;  
DEUTSCHES KREBSFORSCH (DB); POUSTKA ANNEMARIE (DB)

FEATURES  
Source location/Qualifiers  
1..5802  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
107..5464  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB69392.1"  
/translation="MGISTVILEMCLMGQVLTGSGMIPRTDYASLIPSEVPDQV  
AEGPPESTLESTAEAGSIVLSLESTLVARGLIPSESTLAEAGSDGLAL  
RLVNGDRCCGRVEILYRGSGVTCDDWDINDANVCRQLCGMAMAPGNARFGG  
SGPFLADLVRCSGHESYLMSCPHNGMLSHNCGHEDAGVCSAQPQSTLPEPVR  
ISPPVPEGSESSLLRLVNGDRCCGRVEILYRGSGVTCDDWDINDANVCRQLG  
CGMAMAPGNARFGGSGPIVLADVRCSGHESYLMSCPHNGMLTHNCGHEDAGVCS  
APQSRPTSPDPTWTSHTAGPESLLRLVNGDRCCGRVEILYRGSGVTCDDSW  
DTSANVCRQLCGMAMAPGNARFGGSGPIVLADVRCSGHESYLMSCPHNGMLSH  
NCGHEDAGVCSAHSMTSPSPDPTWTSHTAGPESLLRLVNGDRCCGRVE  
VLQSGWTCVCDSDINDANVCRQLCGMAMAPGNARFGGSGPIVLADVRCSGH  
ESYPSCPHNGMLSHNCGHEDAGVCSAQSRLPSPDPTWTSHTAGPESLLRL  
LVNGDRCCGRVEILYRGSGVTCDDWDINDANVCRQLCGMAMAPGNARFGG  
GPIVLADVRCSGHESYLMSCPHNGMLSHNCGHEDAGVCSAQSOPSPDPTWTS  
ASTAGSESSLLRLVNGDRCCGRVEILYRGSGVTCDDWDINDANVCRQLCGMA  
TSAQGNARFGGSGPIVLADVRCSGHESYLMSCPHNGMLSHNCGHEDAGVCSAQS  
OPSPDPTWTSHTAGPESLLRLVNGDRCCGRVEILYRGSGVTCDDWDIND  
ANVCRQLCGMAMAPGNARFGGSGPIVLADVRCSGHESYLMSCPHNGMLSHNCG  
HEDAGVCSAQSOPSPDPTWTSHTAGPESLLRLVNGDRCCGRVEILYRGSGV  
GSGWTCVCDSDINDANVCRQLCGMAMAPGNARFGGSGPIVLADVRCSGHESYL  
MSCPHNGMLTHNCGHEDAGVCSAQSINSTTMMHPTTTTTPSPNCGGLFYLS  
GTSPSPYAYPNNAKCMWIEVNSGRINLGEHNSCPYVYIPGSLN  
SLLGKICNDTRQIFTSSTARKMTHTWTSHTAGPESLLRLVNGDRCCGRVEIL  
SYGACAGVEILYHGWTGTCDDSWTQEARVCRQLCGAVALSAGVSGSGPI  
TLDDVEGSGESTLMQCNRMGFHNCNHEDAGVCSGNHSTPAFLNTRNTDY  
SCGGFLSGPSGDFSPFYGNYPNNACVMDIEVNNYRVTIREDVQLGECNVYDI  
EVFGPYRSSPLRLVCDGARGFTSSSIRPISDHSITRRGFRAYEYSSNS  
TNLLCPHMOASVRSYLSQSGSASDLVSTWNGYECPQTPPNVITTYSCG  
GTFQADNDITDIYNFLTAVSGGIKRTDLRHVSCKMLQNTWMDVYIANTLIHV  
ANNTIQVEEVOYGNFVNISFTYSSSFLYPTSPRYVDLQDLYVQAEIHSADVLP  
LFVDTCASPYSDNFTSLYDLIRSGCYRDYGGYSSPSLRIFRRRAHFLNRP  
SVYLRKMYCRAVDPSSRCYRCGLRKRKDVGSYQEKVDVLPILQGTTPRREER  
R"

BASE COUNT 1253 a 1614 c 1576 g 1359 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 6; Length 5802;  
Best Local Similarity 94.4%; Pred. No. 1.8e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGCGAGCGGATGCG 18  
Db 5435 GGCTGCGAGCTGGATCG 5418

RESULT 36  
AR193170/c  
LOCUS AR193170 5802 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 4 from patent US 6346606.  
ACCESSION AR193170  
VERSION AR193170.1 GI:20239135  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5802)  
AUTHORS Mollenhauer, J. and Poustka, A.  
TITLE Protein containing a scavenger receptor cysteine rich domain  
JOURNAL Patent: US 6346606-A 4 12-FEB-2002;  
DEUTSCHES KREBSFORSCH (DB); POUSTKA ANNEMARIE (DB)

FEATURES  
Source location/Qualifiers  
1..5802  
/organism="unknown"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 1253 a 1614 c 1576 g 1359 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 6; Length 5802;  
Best Local Similarity 94.4%; Pred. No. 1.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGCGAGCGGATGCG 18  
Db 5435 GGCTGCGAGCTGGATCG 5418

RESULT 37  
BD064963/c  
LOCUS BD064963 5802 bp DNA linear PAT 27-AUG-2002  
DEFINITION Protein containing an SRCR domain.  
ACCESSION BD064963  
VERSION BD064963.1 GI:22610566  
KEYWORDS JP 2001509667-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 5802)  
AUTHORS Mollenhauer, J. and Poustka, A.  
TITLE Protein containing an SRCR domain  
JOURNAL Patent: JP 2001509667-A 1 24-JUL-2001;  
DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OFFENTLICHEN RECHTS  
OS Homo sapiens (human)  
PN JP 2001509667-A/1  
PD 24-JUL-2001  
PF 09-JAN-1998 JP 1998530469  
PR 09-JAN-1997 DE 197 00 519.5, 18-JUL-1997 DE 197 30 997.6 PI  
JAN MOLLENHAUER, ANNEMARIE POUSTKA  
PC C12N15/12, C12N15/70, C12N1/21, C12Q1/68, C07K14/47, C07K16/18, PC  
A61K38/17,  
PC A61K48/00, G01N33/50  
CC Protein containing an SRCR domain  
FH key location/Qualifiers  
FT CDS (107)..(5461).

BASE COUNT 1253 a 1614 c 1576 g 1359 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 6; Length 5802;









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 13, 2004, 15:53:18 ; Search time 123.5 Seconds  
(without alignments)  
415.298 Million cell updates/sec

Title: US-09-719-737-9

Sequence: 1 ggcctccgcagcgagatgct 19

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	19	100.0	19	AAZ56193
2	19	100.0	19	ABX12688
3	19	100.0	19	ABX12699
4	19	100.0	19	ABX12700
5	19	100.0	619	AAST0201
6	19	100.0	3394	AAZ51325
7	19	100.0	3439	ABX63546
8	19	100.0	3475	AAQ21453

C	9	19	100.0	21968	25	AAZ51326	Human receptor gen
C	10	19	100.0	86574	24	ABX83560	Human cDNA difere
C	11	18	94.7	19	21	AAZ56194	Oligonucleotide 10
C	12	18	94.7	19	21	AAZ56195	Oligonucleotide 10
C	13	18	94.7	19	25	ABX12689	Human IL-3/IL-5/GM
C	14	18	94.7	19	25	ABX12690	Human IL-3/IL-5/GM
C	15	17	89.5	19	25	ABX12691	Oligonucleotide 10
C	16	17	89.5	19	25	ABX12692	Human IL-3/IL-5/GM
C	17	17	89.5	19	25	ABX12701	Human IL-3/IL-5/GM
C	18	17	89.5	19	25	ABX12702	Human IL-3/IL-5/GM
C	19	16.4	86.3	256	22	AAZ56193	Human immune/haema
C	20	16.4	86.3	436	22	AAZ56194	C. glutamicum SRT
C	21	16.4	86.3	535	22	AAZ56195	Human immune/haema
C	22	16.4	86.3	536	22	AAZ56196	Human immune/haema
C	23	16.4	86.3	825	23	AAZ56197	DNA encoding novel
C	24	16.4	86.3	825	23	AAZ56198	DNA encoding novel
C	25	16.4	86.3	1095	21	AAZ56199	Drosophila odorant
C	26	16.4	86.3	1194	21	AAZ56200	Drosophila melanog
C	27	16.4	86.3	1197	23	ABX12703	Drosophila melanog
C	28	16.4	86.3	1320	24	ABX12704	M. capsulatus gene
C	29	16.4	86.3	1443	24	ABX12705	Human secreted pro
C	30	16.4	86.3	1537	21	AAZ56201	C. glutamicum codin
C	31	16.4	86.3	1824	22	AAZ56202	C. glutamicum SRT
C	32	16.4	86.3	1947	22	AAZ56203	Drosophila melanog
C	33	16.4	86.3	3309	23	AAZ56204	Human SRCR protein
C	34	16.4	86.3	5802	19	AAZ56205	DNA encoding human
C	35	16.4	86.3	8045	22	AAZ56206	DNA encoding human
C	36	16.4	86.3	8047	22	AAZ56207	Protonobacterium
C	37	16.4	86.3	21719	23	AAZ56208	Human SC3 DNA. Ho
C	38	16.4	86.3	28720	19	AAZ56209	C. glutamicum codin
C	39	16.4	86.3	349980	22	AAZ56210	Oligonucleotide 11
C	40	16.4	86.3	19	21	AAZ56211	Human IL-3/IL-5/GM
C	41	16	84.2	19	25	ABX12691	Human IL-3/IL-5/GM
C	42	16	84.2	7217	23	ABX12692	Drosophila melanog
C	43	16	84.2	10131	23	ABX12693	Drosophila melanog
C	44	15.8	83.2	305	21	AAZ56214	Human secreted exp
C	45	15.8	83.2	624	23	AAZ56215	DNA encoding novel

#### ALIGNMENTS

AAZ56193	standard; DNA; 19 BP.
AAZ56193;	
28-MAR-2000	(first entry)
DE	Oligonucleotide 107A for IL-3/IL-5/GM-CSF receptor expression inhibition.
XX	Interleukin-3; IL-3; interleukin-5; IL-5; antisense oligonucleotide;
XX	asbhma; allergy; cancer; receptor expression inhibitor; cytokine;
KW	inflammation; hyperosinophilia; eosinophil proliferation;
KW	granulocyte macrophage colony stimulating factor; GM-CSF; 88.
XX	
OS	Homo sapiens.
XX	
PN	WO9966037-A2.
XX	
PD	23-DEC-1999.
XX	
PF	17-JUN-1999; 99WO-CA00572.
XX	
PR	17-JUN-1998; 98CA-2235420.
XX	
PA	(REBX-) RECH EXPERTISES & DEV MEDICAUX PARENZ IN.
PI	Renzi P;
XX	
DR	WPI; 2000-097743/08.
XX	

PT Antisense oligonucleotides directed to CCR3, interleukin or granulocyte  
PT macrophage colony stimulating factor receptors, used for treating or  
PT preventing asthma, allergies, hypereosinophilia, inflammation or cancer  
XX  
PS Claim 5; Page 23; 72pp; English.  
XX This is an antisense oligonucleotide directed against the common beta  
CC subunit of the interleukin-3 (IL-3) receptor, the interleukin-5 (IL-5)  
CC receptor and the granulocyte macrophage colony stimulating factor  
CC (GM-CSF) receptor. The antisense oligonucleotide inhibits IL-3, IL-5  
CC and GM-CSF receptor expression. IL-3, IL-5 and GM-CSF are cytokines  
CC involved in eosinophil proliferation and survival, they are increased in  
CC asthma and atopic diseases, and are also involved in the indefinite  
CC proliferation of certain neoplastic diseases. The invention relates to  
CC antisense oligonucleotides directed against a nucleic acid sequence  
CC encoding either a chemokine receptor (CCR3), a common subunit of  
CC interleukin-4 (IL-4) and interleukin-13 (IL-13) receptors, or a common  
CC subunit of IL-3, IL-5 and GM-CSF receptors. The antisense  
CC oligonucleotides can be used in the treatment or prevention of asthma,  
CC allergy, hypereosinophilia, general inflammation or cancer.  
XX  
SQ Sequence 19 BP; 2 A; 3 C; 10 G; 4 T; 0 other;  
Query Match 100.0%; Score 19; DB 21; Length 19;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GGGTCTGCAGCGGATGCT 19  
1 ||||||||||||||||  
Db 1 GGGTCTGCAGCGGATGCT 19  
RESULT 2  
ABX12688  
ID ABX12688 standard; DNA; 19 BP.  
XX  
AC ABX12688;  
XX  
DT 10-MAY-2003 (first entry)  
XX  
DE Human IL-3/IL-5/GM-CSF receptor DNA, antisense oligonucleotide #1.  
XX  
KW Human; inflammation; 2',6'-diaminopurine; DAP; antisense therapy;  
KW DAP-modified oligonucleotide; pulmonary disease; respiratory disease;  
KW neurological disease; cardiovascular disease; rheumatological disease;  
KW digestive disease; cutaneous disease; ophthalmological disease;  
KW urinary system disease; pathogen infection; genetic disease; cancer;  
KW airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;  
KW cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;  
KW eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;  
KW hypereosinophilia; cardiant; ophthalmological; cytostatic;  
KW antiasthmatic; antiallergic; antiinflammatory; immunosuppressive;  
KW atopic disease; neoplastic cell proliferation; antisense; IL-3; IL-5;  
KW interleukin-3 receptor; interleukin-5 receptor; GM-CSF; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2003004511-A2.  
XX  
PD 16-JAN-2003.  
XX  
PP 08-JUL-2002; 2002WO-CA01046.  
XX  
PR 06-JUL-2001; 2001US-303071P.  
XX  
PA (TOPI-) TOPIGEN PHARM INC.  
XX  
PI Renzi P, Allam M, Alakhverdi Z;  
XX  
DR WPI, 2003-247944/24.  
XX  
PT Increasing in vivo efficacy of a nucleic acid molecule that is

PT administered to a mammal for inhibiting inflammation in mammals,  
PT involves incorporating into the nucleic acid molecule at least one  
PT nucleotide substitute -  
XX  
PS Claim 28; Page 11; 63pp; English.  
XX  
CC The present invention relates to a method for increasing the in vivo  
CC efficacy of oligonucleotides and inhibiting inflammation. The  
CC oligonucleotides comprise at least one nucleotide substitute of  
CC 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide  
CC substitutions are useful for increasing in vivo efficacy of a nucleic  
CC acid molecule that is administered to a mammal. The DAP-modified  
CC oligonucleotides are useful in antisense therapy for treating and/or  
CC preventing pulmonary/respiratory diseases, neurological diseases,  
CC cardiovascular diseases, rheumatological diseases, digestive diseases,  
CC cutaneous diseases, ophthalmological diseases, urinary system diseases,  
CC pathogen infections, genetic diseases, general inflammation and  
CC cancers. The respiratory system disease is a sickness associated with  
CC an inflammation of the lungs, the airways and/or the nose. The  
CC respiratory system disease is selected from pulmonary fibrosis, adult  
CC respiratory distress syndrome, cystic fibrosis, chronic obstructive  
CC lung disease, chronic bronchitis, eosinophilic bronchitis, asthma,  
CC allergy, allergic rhinitis, sinusitis and hypereosinophilia. The  
CC DAP-modified oligonucleotides are more stable in the body, more  
CC effective, and less toxic than standard antisense oligonucleotides.  
CC DAP or its analogues are more effective than other substitutes of  
CC adenosine. ABX12681-ABX12698 represent antisense oligonucleotides  
CC for treating or preventing atopic diseases and neoplastic cell  
CC proliferation.  
XX  
SQ Sequence 19 BP; 2 A; 3 C; 10 G; 4 T; 0 other;  
Query Match 100.0%; Score 19; DB 25; Length 19;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GGGTCTGCAGCGGATGCT 19  
1 ||||||||||||||||  
Db 1 GGGTCTGCAGCGGATGCT 19  
RESULT 3  
ABX12699/c  
ID ABX12699 standard; DNA; 19 BP.  
XX  
AC ABX12699;  
XX  
DT 10-MAY-2003 (first entry)  
XX  
DE Human IL-3/IL-5/GM-CSF receptor DNA, sense oligonucleotide 107S.  
XX  
KW Human; inflammation; 2',6'-diaminopurine; DAP; antisense therapy;  
KW DAP-modified oligonucleotide; pulmonary disease; respiratory disease;  
KW neurological disease; cardiovascular disease; rheumatological disease;  
KW digestive disease; cutaneous disease; ophthalmological disease;  
KW urinary system disease; pathogen infection; genetic disease; cancer;  
KW airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;  
KW cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;  
KW eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;  
KW hypereosinophilia; cardiant; ophthalmological; cytostatic;  
KW antiasthmatic; antiallergic; antiinflammatory; immunosuppressive;  
KW atopic disease; neoplastic cell proliferation; IL-3; IL-5;  
KW interleukin-3 receptor; interleukin-5 receptor; GM-CSF; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2003004511-A2.  
XX  
PD 16-JAN-2003.  
XX  
PP 08-JUL-2002; 2002WO-CA01046.  
XX  
PR 06-JUL-2001; 2001US-303071P.  
XX  
PT

```
XX (TOPI-) TOPIGEN PHARM INC.
XX
XX Renzi P, Allam M, Allakhverdi Z;
XX WPI; 2003-247944/24.
XX
XX Increasing in vivo efficacy of a nucleic acid molecule that is
XX administered to a mammal for inhibiting inflammation in mammals,
XX involves incorporating into the nucleic acid molecule at least one
XX nucleotide substitute
XX
XX Examples; Page 18; 63pp; English.
XX
XX The present invention relates to a method for increasing the in vivo
XX efficacy of oligonucleotides and inhibiting inflammation. The
XX oligonucleotides comprise at least one nucleotide substitute of
XX 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide
XX substitutions are useful for increasing in vivo efficacy of a nucleic
XX acid molecule that is administered to a mammal. The DAP-modified
XX oligonucleotides are useful in antisense therapy for treating and/or
XX preventing pulmonary/respiratory diseases, neurological diseases,
XX cardiovascular diseases, rheumatological diseases, digestive diseases,
XX cutaneous diseases, ophthalmological diseases, urinary system diseases,
XX pathogen infections, genetic diseases, general inflammation and
XX cancers. The respiratory system disease is a sickness associated with
XX an inflammation of the lungs, the airways and/or the nose. The
XX respiratory system disease is selected from pulmonary fibrosis, adult
XX respiratory distress syndrome, cystic fibrosis, chronic obstructive
XX lung disease, chronic bronchitis, eosinophilic bronchitis, asthma,
XX allergy, allergic rhinitis, sinusitis and hyperesinophilia. The
XX DAP-modified oligonucleotides are more stable in the body, more
XX effective, and less toxic than standard antisense oligonucleotides.
XX DAP or its analogues are more effective than other substitutes of
XX adenosine. The present sequence represents a sense oligonucleotide
XX used in the examples of the present invention.
XX
XX Sequence 19 BP; 4 A; 10 C; 3 G; 2 T; 0 other;
XX
XX Query Match 100.0%; Score 19; DB 25; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 10;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGGTCTGCAGCGGATGCT 19
XX |||||
XX 19 GGGTCTGCAGCGGATGCT 1
XX
XX RESULT 4
XX ID ABX12700 standard; DNA; 19 BP.
XX AC ABX12700;
XX DT 10-MAY-2003 (first entry)
XX
XX Human IL-3/IL-5/GM-CSF receptor DNA, antisense oligonucleotide 107A.
XX
XX Human; inflammation; 2',6'-diaminopurine; DAP; antisense therapy;
XX DAP-modified oligonucleotide; pulmonary disease; respiratory disease;
XX neurological disease; cardiovascular disease; rheumatological disease;
XX digestive disease; cutaneous disease; ophthalmological disease;
XX urinary system disease; pathogen infection; genetic disease; cancer;
XX airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;
XX cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;
XX eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;
XX hyperesinophilia; cardiant; ophthalmological; cyclostatic;
XX antiaesthetic; antiinflammatory; immunosuppressive;
XX atopic disease; neoplastic cell proliferation; antisense; IL-3; IL-5;
XX interleukin-3 receptor; interleukin-5 receptor; GM-CSF; ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX
```

```
PN WO2003004511-A2.
XX
XX 16-JAN-2003.
XX
XX 08-JUL-2002; 2002WO-CA01046.
XX
XX 06-JUL-2001; 2001US-303071P.
XX
XX (TOPI-) TOPIGEN PHARM INC.
XX
XX Renzi P, Allam M, Allakhverdi Z;
XX WPI; 2003-247944/24.
XX
XX Increasing in vivo efficacy of a nucleic acid molecule that is
XX administered to a mammal for inhibiting inflammation in mammals,
XX involves incorporating into the nucleic acid molecule at least one
XX nucleotide substitute
XX
XX Examples; Page 18; 63pp; English.
XX
XX The present invention relates to a method for increasing the in vivo
XX efficacy of oligonucleotides and inhibiting inflammation. The
XX oligonucleotides comprise at least one nucleotide substitute of
XX 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide
XX substitutions are useful for increasing in vivo efficacy of a nucleic
XX acid molecule that is administered to a mammal. The DAP-modified
XX oligonucleotides are useful in antisense therapy for treating and/or
XX preventing pulmonary/respiratory diseases, neurological diseases,
XX cardiovascular diseases, rheumatological diseases, digestive diseases,
XX cutaneous diseases, ophthalmological diseases, urinary system diseases,
XX pathogen infections, genetic diseases, general inflammation and
XX cancers. The respiratory system disease is a sickness associated with
XX an inflammation of the lungs, the airways and/or the nose. The
XX respiratory system disease is selected from pulmonary fibrosis, adult
XX respiratory distress syndrome, cystic fibrosis, chronic obstructive
XX lung disease, chronic bronchitis, eosinophilic bronchitis, asthma,
XX allergy, allergic rhinitis, sinusitis and hyperesinophilia. The
XX DAP-modified oligonucleotides are more stable in the body, more
XX effective, and less toxic than standard antisense oligonucleotides.
XX DAP or its analogues are more effective than other substitutes of
XX adenosine. The present sequence represents an antisense oligonucleotide
XX used in the examples of the present invention.
XX
XX Sequence 19 BP; 2 A; 3 C; 10 G; 4 T; 0 other;
XX
XX Query Match 100.0%; Score 19; DB 25; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 10;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGGTCTGCAGCGGATGCT 19
XX |||||
XX 1 GGGTCTGCAGCGGATGCT 19
XX
XX RESULT 5
XX ID AAS70201/C
XX AC AAS70201;
XX DT 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #6005.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO200175067-A2.
XX
XX PN
XX
XX 11-OCT-2001.
XX
```

XX 30-MAR-2001; 2001WO-US08631.  
PF  
XX  
XX 31-MAR-2000; 2000US-0540217.  
PR  
XX 23-AUG-2000; 2000US-0649167.  
PR  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Drmanac RT, Liu C, Tang YT;  
PI  
XX  
XX WPI; 2001-639362/73.  
DR  
XX P-PSDB; ABG06014.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnosis, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX  
XX Claim 1; SEQ ID No 6005; 103bp; English.  
PS  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations in  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 619 BP; 121 A; 203 C; 182 G; 113 T; 0 other;  
XX  
XX  
XX Query Match 100.0%; Score 19; DB 23; Length 619;  
XX Best Local Similarity 100.0%; Pred. No. 12;  
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX  
XX 1 GGGTCTGCAGCGGATGCT 19  
QY |||||  
DB 336 GGGTCTGCAGCGGATGCT 318  
XX  
XX  
XX RESULT 6  
XX AAD51325/c  
XX ID AAD51325 standard; cDNA; 3394 BP.  
XX  
XX AAD51325;  
AC  
XX  
XX 16-APR-2003 (first entry)  
DT  
XX  
XX Human receptor cDNA.  
DE  
XX  
XX Human; receptor; leukaemia; cancer; gene therapy; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX 5'UTR 1..188  
XX /\*tag= a  
XX CDS 189..2273  
XX /\*tag= b  
XX /product= "Human receptor"  
XX  
XX  
XX

FT 3'UTR 2274..3394  
/\*tag= c  
XX  
XX WO200292835-A2.  
XX  
XX  
XX 21-NOV-2002.  
PD  
XX  
XX 07-MAY-2002; 2002WO-US14274.  
XX  
XX  
XX 10-MAY-2001; 2001US-0851985.  
PR  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX  
XX Gong F, Yan C, Di Francesco V, Beasley EM;  
PI WPI; 2003-129304/12.  
XX  
XX P-PSDB; AAE33203.  
DR  
XX  
XX  
XX New human receptor proteins, useful for preparing a composition for  
PT treating a disease or condition mediated by a human receptor protein  
PT e.g., leukemia or cancer -  
XX  
XX  
XX Claim 4; Fig 1A; 39pp; English.  
PS  
XX  
XX The invention relates to novel human receptor proteins and nucleic acid  
CC molecules encoding such proteins. The invention is useful for preparing  
CC a composition for treating a disease or condition mediated by a human  
CC receptor protein e.g. leukemia or cancer. The invention is also useful  
CC in gene therapy. The present sequence is human receptor cDNA.  
XX  
XX  
SQ Sequence 3394 BP; 713 A; 1011 C; 927 G; 743 T; 0 other;  
XX  
XX  
XX Query Match 100.0%; Score 19; DB 25; Length 3394;  
XX Best Local Similarity 100.0%; Pred. No. 14;  
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX  
XX 1 GGGTCTGCAGCGGATGCT 19  
QY |||||  
DB 285 GGGTCTGCAGCGGATGCT 267  
XX  
XX  
XX RESULT 7  
XX ABX63546/c  
XX ID ABX63546 standard; cDNA; 3439 BP.  
XX  
XX  
XX ABX63546;  
AC  
XX  
XX 26-FEB-2003 (first entry)  
DT  
XX  
XX Human cDNA #546 differentially expressed in activated vascular tissue.  
DE  
XX  
XX Human; gene; ss; vascular tissue; cytosolic; atherosclerosis;  
XX cardiac; hypotensive; antidiabetic; gynaecological; vasotropic;  
XX cerebroprotective; gene therapy; vascular disease; cancer; coronary;  
XX artery disease; hypertension; diabetes; pre-eclampsia; restenosis;  
XX ischaemia-reperfusion injury; stroke;  
XX  
XX  
XX Homo sapiens.  
XX  
XX  
XX US2002137081-A1.  
XX  
XX 26-SEP-2002.  
PD  
XX  
XX 08-JAN-2002; 2002US-0044090.  
PF  
XX  
XX 28-JUL-2000; 2000US-222469P.  
PR  
XX 08-JAN-2001; 2001US-260483P.  
PR  
XX  
XX (BAND/) BANDMAN O.  
XX  
XX Bandman O;  
XX  
XX WPI; 2003-110597/10.  
XX  
XX  
XX

XX Combination for diagnosing, staging, treating, or monitoring the  
PT progression of treatment of a vascular disease, e.g. atherosclerosis,  
PT comprises several cDNAs that are differentially expressed in activated  
PT vascular tissue -

XX Claim 1; Page -; 18pp; English.

XX This invention relates to a combination comprising several cDNAs that  
CC are differentially expressed in activated vascular tissue. The invention  
CC also discloses a high throughput method for detecting differentially  
CC expressed cDNAs in a sample. The cDNAs of the invention may have  
CC antiatherosclerotic, cytostatic, cardiant, hypotensive, antidiabetic,  
CC gynaecological, vasotropic and cerebroprotective activities and may be  
CC used in gene therapy. The cDNAs of the invention may be used in a  
CC high-throughput methods for detecting differential expression of one or  
CC more cDNAs in a sample, or screening several molecules or compounds to  
CC identify a molecule or compound that specifically binds a cDNA of the  
CC invention. A protein encoded by the cDNA may be used to screen several  
CC molecules or compounds to identify a ligand that specifically binds to  
CC the protein, or to produce or purify an antibody to the protein that can  
CC be used to detect a protein in a sample or purify a natural or  
CC recombinant protein from a sample. The nucleotides may be useful for  
CC diagnosing, staging, treating, or monitoring the progression of  
CC treatment of a vascular disease, e.g. atherosclerosis, cancer, coronary  
CC artery disease, hypertension, diabetes, pre-eclampsia, ischaemia-  
CC reperfusion injury, restenosis, or stroke. The cDNAs can also be used  
CC for large-scale genetic or gene expression analysis of several new  
CC nucleic acid molecules. Antibodies to the proteins encoded by the  
CC cDNAs are useful for diagnosing pre-pathologic disorders, and chronic  
CC or acute diseases associated with abnormalities in the expression,  
CC amount or distribution of the protein. The present sequence  
CC represents a cDNA of the invention that is differentially expressed in  
CC activated vascular tissue.  
CC Note: The sequence data for this patent did not form part of the  
CC specification, but was obtained in electronic format directly from USPTO  
CC at <http://seqdata.uspto.gov/sequence.html?docid=20020137081>.

XX Sequence 3439 BP; 675 A; 1122 C; 999 G; 643 T; 0 other;

Query Match 100.0%; Score 19; DB 25; Length 3439;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTCTGACGGCGATGCT 19  
|||  
DB 273 GGGTCTGACGGCGATGCT 255

RESULT 8  
AAQ21453/C  
ID AAQ21453 standard; DNA; 3475 BP.

XX AAQ21453;

XX 17-MAY-1992 (first entry)

XX Sequence encoding beta-chain of a human granulocyte-macrophage  
DE colony stimulating factor (GM-CSF) receptor.

XX Agonist; antagonist; myeloid leukaemia; therapy; screening;  
KM diagnosis; granulocyte-macrophage colony stimulating factor; ss  
XX Homo sapiens.

XX Key Location/Qualifiers  
FH CDS 201..2846  
FT /\*tag= a

XX W09201788-A.  
XX 06-FEB-1992.

PF 16-JUL-1991; 91WO-US04846.  
XX 18-JUL-1990; 90US-0554745.  
XX (SCHE ) SCHERING CORP.

XX Hayashida K, Kitamura T, Miyajima A;

XX WPI; 1992-064947/08.  
DR P-FSDB; AAR20982.

XX Beta-chain of human granulocyte-macrophage CSF receptor - used  
PT for screening agonists and antagonists of human GM-CSF, e.g. for  
PT diagnosing myeloid leukaemia

XX Claim 6; Page 13-14; 26pp; English.

XX The nucleic acid encoding the human GM-CSF beta-chain is isolated  
CC from a cDNA library prep. from poly(A)+ RNA from TF-1 cells. The  
CC high affinity human GM-CSF receptor (Kd < 1 nM) can be used for  
CC screening candidate GM-CSF agonists and antagonists e.g. for  
CC treating myeloid leukaemias.

XX Sequence 3475 BP; 685 A; 1135 C; 1001 G; 654 T; 0 other;

Query Match 100.0%; Score 19; DB 13; Length 3475;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTCTGACGGCGATGCT 19  
|||  
DB 297 GGGTCTGACGGCGATGCT 279

RESULT 9  
AADS1326/C  
ID AADS1326 standard; DNA; 21968 BP.

XX AADS1326;

XX 16-APR-2003 (first entry)

XX Human receptor gene.

XX Human; receptor; leukaemia; cancer; gene therapy; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH variation replace (183, T)

FT /\*tag= a /standard\_name= "Single nucleotide polymorphism"

FT replace (312, T) /tag= b

FT /\*tag= b /standard\_name= "Single nucleotide polymorphism"

FT replace (400, C) /tag= c

FT /\*tag= c /standard\_name= "Single nucleotide polymorphism"

FT replace (1449, A, G) /tag= d

FT /\*tag= d /standard\_name= "Single nucleotide polymorphism"

FT replace (1453, A) /tag= e

FT /\*tag= e /standard\_name= "Single nucleotide polymorphism"

FT CDS 2176..19200 /tag= f

FT /\*tag= f /product= "Human receptor"

FT exon 2176..2251 /tag= g

FT intron 2252..3211 /tag= h

FT variation replace (2441, C) /tag= i

```

FT /standard_name= "Single nucleotide polymorphism"
FT replace (2705, C)
FT /tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT replace (2935, A)
FT /tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT replace (3007, G)
FT /tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT 3212..3335
FT /tag= m
FT 3336..5951
FT /tag= n
FT replace (5705, C)
FT /tag= o
FT /standard_name= "Single nucleotide polymorphism"
FT replace (5783, G)
FT /tag= p
FT /standard_name= "Single nucleotide polymorphism"
FT 5952..6142
FT /tag= q
FT 6143..9366
FT /tag= r
FT replace (6905, -)
FT /tag= s
FT /standard_name= "Single nucleotide polymorphism"
FT replace (7443, -)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT replace (7445, -)
FT /tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT replace (7561, T)
FT /tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT replace (7911, C)
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT 9367..9524
FT /tag= x
FT replace (9427, -)
FT /tag= y
FT /standard_name= "Single nucleotide polymorphism"
FT 9525..9603
FT /tag= z
FT 9604..9772
FT /tag= aa
FT 9773..10339
FT /tag= ab
FT 10340..10475
FT /tag= ac
FT 10476..10637
FT /tag= ad
FT 10638..10795
FT /tag= ae
FT 10796..12729
FT /tag= af
FT 12730..12869
FT /tag= ag
FT 12870..13796
FT /tag= ah
FT replace (13559, T)
FT /tag= ai
FT /standard_name= "Single nucleotide polymorphism"
FT 13797..13959
FT /tag= aj
FT replace (13922, C)
FT /tag= ak
FT /standard_name= "Single nucleotide polymorphism"
FT 13960..15315
FT /tag= al
FT replace (14005, C)

```

```

FT /tag= am
FT /standard_name= "Single nucleotide polymorphism"
FT replace (14702, C)
FT /tag= an
FT /standard_name= "Single nucleotide polymorphism"
FT replace (15228, G)
FT /tag= ao
FT /standard_name= "Single nucleotide polymorphism"
FT 15316..15406
FT /tag= ap
FT 15407..15595
FT /tag= aq
FT 15596..15653
FT /tag= ar
FT 15654..16515
FT /tag= as
FT 16516..16619
FT /tag= at
FT 16620..17343
FT /tag= au
FT 17344..17513
FT /tag= av
FT 17514..18853
FT /tag= aw
FT 18854..19200
FT /tag= ax
FT replace (19251, C)
FT /tag= ay
FT /standard_name= "Single nucleotide polymorphism"
FT replace (21335, C)
FT /tag= az
FT /standard_name= "Single nucleotide polymorphism"
FT replace (21698, G)
FT /tag= ba
FT /standard_name= "Single nucleotide polymorphism"

```

```

WO200292835-A2.
21-NOV-2002.
07-MAY-2002; 2002WO-US14274.
10-MAY-2001; 2001US-0851985.
(PKKE ) PE CORP NY.
Gong F, Yan C, Di Francesco V, Beasley EM;
WPI, 2003-129304/12.
P-PsDB; AAE33203.
New human receptor proteins, useful for preparing a composition for
treating a disease or condition mediated by a human receptor protein
e.g., leukemia or cancer -
PS
Claim 4: Fig 3; 39pp; English.
The invention relates to novel human receptor proteins and nucleic acid
molecules encoding such proteins. The invention is useful for preparing
a composition for treating a disease or condition mediated by a human
receptor protein e.g. leukemia or cancer. The invention is also useful
in gene therapy. The present sequence is human receptor gene.
XX

```

Sequence 21968 BP; 4659 A; 5655 C; 5769 G; 5017 T; 868 other;

Query Match 100.0%; Score 19; DB 25; Length 21968;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGGCTGACGCGGATGCT 19
DB 3232 GGGCTGACGCGGATGCT 3214

```



RESULT 10  
ABK83560/c  
ID ABK83560 standard; CDNA; 86574 BP.  
XX  
AC  
XX ABK83560;  
XX  
DT 14-AUG-2002 (first entry)  
XX  
DE Human CDNA differentially expressed in granulocytic cells #131.  
XX  
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
KW viral infection; parasitic infection; protozoal infection;  
KW fungal infection; sterile inflammatory disease; psoriasis;  
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
KW adult respiratory distress syndrome; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; periodontal disease;  
KW granulocyte activation; chronic inflammation; allergy.  
XX  
OS Homo sapiens.  
XX  
PN WO200228999-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 03-OCT-2001; 2001WO-US30821.  
XX  
PR 03-OCT-2000; 2000US-237189P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
XX  
DR WPI; 2002-435328/46.  
XX  
PT Detecting granulocyte activation by detecting differential expression  
PT of genes associated with granulocyte activation, which serves as  
PT diagnostic markers that is useful for monitoring disease states and  
PT drug toxicity -  
XX  
PS Claim 1; SEQ ID No 131; 114pp; English.  
XX  
CC The invention relates to detecting (M1) granulocyte (GC) activation  
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
CC DNA chip analysis as given in the specification, and comparing  
CC the expression level to an expression level in an unactivated  
CC GC, where differential expression of Gs is indicative of GCA.  
CC Also included are modulating (M2) GA by contacting GC with an agent  
CC that alters the expression of at least one gene in Gs; (2) screening (M3)  
CC for an agent capable of modulating GCA or an inflammation (especially  
CC chronic) in a tissue, an allergic response in a subject, exposure of a  
CC subject to a pathogen or sterile inflammatory disease using the  
CC gene expression profile; (3) detecting (M4) an inflammation (especially  
CC chronic) in a tissue, an allergic response in a subject, exposure of a  
CC subject to a pathogen or sterile inflammatory disease, by detecting the  
CC level of expression in a sample of the tissue of gene(s) from Gs, where  
CC the level of expression of the gene is indicative of inflammation;  
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,  
CC an allergic response in a subject, exposure of a subject to a pathogen  
CC or sterile inflammatory disease, by contacting a tissue having  
CC inflammation with an agent that modulates the expression of gene(s)  
CC from Gs in the tissue. M1 is useful for detecting GCA, M2 is useful for  
CC modulating GCA, M3 is useful for screening an agent capable of modulating  
CC GCA preferably in an inflammation in a tissue; M4 is useful for  
CC detecting an inflammation (especially chronic) in a tissue, an allergic  
CC response in a subject, exposure of a subject to a pathogen or sterile  
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,  
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal  
CC reperfusion injury, ARDS, adult respiratory distress syndrome,  
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,  
CC periodontal disease, also bacterial infection, viral infection,  
CC parasitic infection, protozoal infection, fungal infection and M5 is

CC useful for treating one of the above conditions. The present  
CC sequence represents a gene differentially expressed in granulocytes.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 86574 BP; 22071 A; 20398 C; 21552 G; 22553 T; 0 other;  
XX  
Query Match 100.0%; Score 19; DB 24; Length 86574;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
Oy 1 GGGTCGACGGCGATGCT 19  
Db 74634 GGGTCGACGGCGATGCT 74616  
XX  
RESULT 11  
AAZ56194  
ID AAZ56194 standard; DNA; 19 BP.  
XX  
AC AAZ56194;  
XX  
DT 28-MAR-2000 (first entry)  
XX  
DE Oligonucleotide 106 for IL-3/IL-5/GM-CSF receptor expression inhibition.  
XX  
KW Interleukin-3; IL-3; interleukin-5; IL-5; antisense oligonucleotide;  
KW asthma; allergy; cancer; receptor expression inhibitor; cytokine;  
KW inflammation; hypersensitivity; eosinophil proliferation;  
KW granulocyte macrophage colony stimulating factor; GM-CSF; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9966037-A2.  
XX  
PD 23-DEC-1999.  
XX  
PF 17-JUN-1999; 99WO-CA00572.  
XX  
PR 17-JUN-1998; 98CA-2235420.  
XX  
PA (REX-) RECH EXPERTISES & DEV MEDICADUX PARENZ IN.  
XX  
PI Renzi P;  
XX  
DR WPI; 2000-097743/08.  
XX  
PT Antisense oligonucleotides directed to CCR3, interleukin or granulocyte  
PT macrophage colony stimulating factor receptors, used for treating or  
PT preventing asthma, allergies, hypersensitivity, inflammation or cancer  
PT -  
XX  
PS Claim 5; Page 25; 72pp; English.  
XX  
CC This is an antisense oligonucleotide directed against the common beta  
CC subunit of the interleukin-3 (IL-3) receptor, the interleukin-5 (IL-5)  
CC receptor and the granulocyte macrophage colony stimulating factor  
CC (GM-CSF) receptor. The antisense oligonucleotide inhibits IL-3, IL-5  
CC and GM-CSF receptor expression. IL-3, IL-5 and GM-CSF are cytokines  
CC involved in eosinophil proliferation and survival, they are increased in  
CC asthma and atopic diseases, and are also involved in the indefinite  
CC proliferation of certain neoplastic diseases. The invention relates to  
CC antisense oligonucleotides directed against a nucleic acid sequence  
CC encoding either a chemokine receptor (CCR3), a common subunit of  
CC interleukin-4 (IL-4) and interleukin-13 (IL-13) receptors, or a common  
CC subunit of IL-3, IL-5 and GM-CSF receptors. The antisense  
CC oligonucleotides can be used in the treatment or prevention of asthma,  
CC allergy, hypersensitivity, general inflammation or cancer.  
XX  
SQ Sequence 19 BP; 2 A; 3 C; 9 G; 5 T; 0 other;

Query Match 94.7%; Score 18; DB 21; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GGCTGACGCGGATGT 19  
 DB 1 GGCTGACGCGGATGT 18

## RESULT 12

AAZ56195  
 ID AAZ56195 standard; DNA; 19 BP.

AC AAZ56195;

DT 28-MAR-2000 (first entry)

DE Oligonucleotide 108 for IL-3/IL-5/GM-CSF receptor expression inhibition.

KM Interleukin-3; IL-3; interleukin-5; IL-5; antisense oligonucleotide;

KM asthma; allergy; cancer; receptor expression inhibitor; cytokine;

KM inflammation; hypereosinophilia; eosinophil proliferation;

KM granulocyte macrophage colony stimulating factor; GM-CSF; ss.

OS Homo sapiens.

PM WO966037-A2.

PD 23-DEC-1999.

PF 17-JUN-1999; 99WO-CA00572.

PR 17-JUN-1998; 98CA-2235420.

PA (REEX-) RECH EXPERTISES & DEV MEDICAUX PARENZ IN.

PI Renzi P;

DR WPI: 2000-097743/08.

PT Antisense oligonucleotides directed to CCR3, interleukin or granulocyte  
 macrophage colony stimulating factor receptors, used for treating or  
 preventing asthma, allergies, hypereosinophilia, inflammation or cancer

PS Claim 5; Page 25; 72pp; English.

CC This is an antisense oligonucleotide directed against the common beta  
 CC subunit of the interleukin-3 (IL-3) receptor, the interleukin-5 (IL-5)  
 CC receptor, and the granulocyte macrophage colony stimulating factor  
 CC (GM-CSF) receptor. The antisense oligonucleotide inhibits IL-3, IL-5  
 CC and GM-CSF receptor expression. IL-3, IL-5 and GM-CSF are cytokines  
 CC involved in eosinophil proliferation and survival, they are increased in  
 CC asthma and atopic diseases, and are also involved in the indefinite  
 CC proliferation of certain neoplastic diseases. The invention relates to  
 CC antisense oligonucleotides directed against a nucleic acid sequence  
 CC encoding either a chemokine receptor (CCR3), a common subunit of  
 CC interleukin-4 (IL-4) and interleukin-13 (IL-13) receptors, or a common  
 CC subunit of IL-3, IL-5 and GM-CSF receptors. The antisense  
 CC oligonucleotides can be used in the treatment or prevention of asthma,  
 CC allergy, hypereosinophilia, general inflammation or cancer.

SO Sequence 19 BP; 3 A; 3 C; 10 G; 3 T; 0 other;

Query Match 94.7%; Score 18; DB 21; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGCTGACGCGGATGT 18  
 DB 2 GGCTGACGCGGATGT 19

## RESULT 13

ID ABX12689  
 ID ABX12689 standard; DNA; 19 BP.

AC ABX12689;

DT 10-MAY-2003 (first entry)

DE Human IL-3/IL-5/GM-CSF receptor DNA, antisense oligonucleotide #2.

KM Human; inflammation; 2',6'-diaminopurine; DAP; antisense therapy;

KM DAP-modified oligonucleotide; pulmonary disease; respiratory disease;

KM neurological disease; cardiovascular disease; rheumatological disease;

KM digestive disease; cutaneous disease; ophthalmological disease;

KM urinary system disease; pathogen infection; genetic disease; cancer;

KM airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;

KM cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;

KM eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;

KM hypereosinophilia; caridant; ophthalmological; cycostatic;

KM antiasthmatic; anti-allergic; anti-inflammatory; immunosuppressive;

KM atopic disease; neoplastic cell proliferation; antisense; IL-3; IL-5;

KM interleukin-3 receptor; interleukin-5 receptor; GM-CSF; ss.

OS Homo sapiens.

PM WO2003004511-A2.

PD 16-JAN-2003.

PF 08-JUL-2002; 2002WO-CA01046.

PR 06-JUL-2001; 2001US-303071P.

PA (TOPIC-) TOPIGEN PHARM INC.

PI Renzi P, Allam M, Allakherdai Z;

DR WPI: 2003-247944/24.

PT Increasing in vivo efficacy of a nucleic acid molecule that is  
 PT administered to a mammal for inhibiting inflammation in mammals.  
 PT involves incorporating into the nucleic acid molecule at least one  
 PT nucleotide substitute

PS Claim 28; Page 11; 63pp; English.

CC The present invention relates to a method for increasing the in vivo  
 CC efficacy of oligonucleotides and inhibiting inflammation. The  
 CC oligonucleotides comprise at least one nucleotide substitute of  
 CC 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide  
 CC substitutions are useful for increasing in vivo efficacy of a nucleic  
 CC acid molecule that is administered to a mammal. The DAP-modified  
 CC oligonucleotides are useful in antisense therapy for treating and/or  
 CC preventing pulmonary/respiratory disease, neurological diseases,  
 CC cardiovascular diseases, rheumatological diseases, digestive diseases,  
 CC cutaneous diseases, ophthalmological diseases, urinary system diseases,  
 CC pathogen infections, genetic diseases, general inflammation and  
 CC cancers. The respiratory system disease is a sickness associated with  
 CC an inflammation of the lungs, the airways and/or the nose. The  
 CC respiratory system disease is selected from pulmonary fibrosis, adult  
 CC respiratory distress syndrome, cystic fibrosis, chronic obstructive  
 CC lung disease, chronic bronchitis, eosinophilic bronchitis, asthma,  
 CC allergy, allergic rhinitis, sinusitis and hypereosinophilia. The  
 CC DAP-modified oligonucleotides are more stable in the body, more  
 CC effective, and less toxic than standard antisense oligonucleotides.  
 CC DAP or its analogues are more effective than other substitutes of  
 CC adenosine. ABX12681-ABX12698 represent antisense oligonucleotides  
 CC for treating or preventing atopic diseases and neoplastic cell  
 CC proliferation.

SO Sequence 19 BP; 2 A; 3 C; 9 G; 5 T; 0 other;

Query Match 94.7%; Score 18; DB 25; Length 19;

Best Local Similarity 100.0%; Pred. No. 30;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Matches	18:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
2000 20000 200000 1000000, FREQ: NO: 50,									

OY		2	GCTCTGCAGCGGATGGT	19
Dd		1	GCTCTGCAGCGGATGGT	18

RESULT 14  
ABX12690  
ID ABX12690 standard; DNA; 19 BP.

AC	ABX12690;
XX	
DT	10-MAY-2003 (first entry)

DE Human IL-3/IL-5/GM-CSF receptor DNA, antisense oligonucleotide #3.

KM Human, inflammation; 2,6'-diaminopurine; DAP; anti-tense therapy;  
KM DAP-modified oligonucleotide; pulmonary disease; respiratory disease;  
KM neurological disease; cardiovascular disease; rheumatological disease;  
KM digestive disease; cutaneous disease; ophthalmological disease;  
KM urinary system disease; pathogen infection; genetic disease; cancer;  
KM airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;  
KM cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;  
KM eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;  
KM hypersensitivity; cardiant; ophthalmological; cytostatic;  
KM antiaesthetic; anti-allergic; anti-inflammatory; immunosuppressive;  
KM atopic disease; neoplastic cell proliferation; anti-tense; IL-3; IL-5;  
KM interleukin-3 receptor; interleukin-5 receptor; GM-CSF; ss.

**OS Homo sapiens.**

PN WO2003004511-A2.

PD 16-JAN-2003.

PF 08-JUL-2002; 2002WO-CA01046.

PR 06-JUL-2001; 2001US-303071P.

PA (TOPI-) TOPIGEN PHARM INC.

PI Renzi P, Alam M, Alakhverdi Z;

DR WPI; 2003-247944/24.

PT Increasing in vivo efficacy of a nucleic acid molecule that is administered to a mammal for inhibiting inflammation in mammals, PT involves incorporating into the nucleic acid molecule at least one PT nucleotide substitute -

PS Claim 28; Page 11; 63pp; English.

The present invention relates to a method for increasing the in vivo efficacy of oligonucleotides and inhibiting inflammation. The oligonucleotides comprise at least one nucleotide substitute of 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide substitutions are useful for increasing in vivo efficacy of a nucleic acid molecule that is administered to a mammal. The DAP-modified oligonucleotides are useful in antisense therapy for treating and/or preventing pulmonary/respiratory diseases, neurological diseases, cardiovascular diseases, rheumatological diseases, digestive diseases, cutaneous diseases, ophthalmological diseases, urinary system diseases, pathogen infections, genetic diseases, general inflammation and cancers. The respiratory system disease is a sickness associated with an inflammation of the lungs, the airways and/or the nose. The respiratory system disease is selected from pulmonary fibrosis, adult respiratory distress syndrome, cystic fibrosis, chronic obstructive lung disease, chronic bronchitis, eosinophilic bronchitis, asthma, allergy, allergic rhinitis, sinusitis and hyperosinophilia. The DAP-modified oligonucleotides are more stable in the body, more effective, and less toxic than standard antisense oligonucleotides.

CC DMP or its analogues are more effective than other substitutes of  
CC adenosine. ABX1261-ABX1268 represent antisense oligonucleotides  
CC for treating or preventing atopic diseases and neoplastic cell  
CC proliferation.

Sequence 19 BP; 3 A; 3 C; 10 G; 3 T; 0 other;

Query Match	94.7%;	Score 18;	DB 25;	Length 19;
Best Local Similarity	100.0%;	Pred. No. 30;		
Matches 18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GGGCTCTGACGGGATGG	18
Db	2	GGGCTCTGACGGGATGG	19

RESULT 15  
AAZ56199  
ID AAZ56199 standard; DNA; 19 BP.

AC AAZ56199;

DT 28-MAR-2000 (first entry)

DE Oligonucleotide 105 for IL-3/IL-5/GM-CSF receptor expression inhibition.

KM Interleukin-3, IL-3, interleukin-5, IL-5, antisense oligonucleotide  
KM aetna; allergy; cancer; receptor expression inhibitor; cytokine;  
KM inflammation; hypereosinophilia; eosinophil proliferation;  
KM granulocyte macrophage colony stimulating factor; GM-CSF, ss.

**OS Homo sapiens.**

PN WO9966037-A2.

PD 23-DEC-1999

PF 17-JUN-1999; 99WO-CA00572.

PR 17-JUN-1998; 98CA-2235420.

PA (REEX-) RECH EXPERTISES & DEV MEDICAUX PARENZ IN.

PI Renzi P;

DR WPI; 2000-097743/08.

PT Antisense oligonucleotides directed to CCR3, interleukin or granulocyte  
PT macrophage colony stimulating factor receptors, used for treating or  
PT preventing asthma, allergies, hypersosinophilia, inflammation or cancer  
PT -

PS Claim 5; Page 25; 72pp; English.

CC This is antisense oligonucleotide directed against the interleukin beta  
CC subunit of the interleukin-3 (IL-3) receptor, the interleukin-5 (IL-5)  
CC receptor and the granulocyte macrophage colony stimulating factor  
CC (GM-CSF) receptor. The antisense oligonucleotide inhibits IL-3, IL-5  
CC and GM-CSF receptor expression. IL-3, IL-5 and GM-CSF are cytokines  
CC involved in eosinophil proliferation and survival, they are increased  
CC in asthma and atopic diseases, and are also involved in the indefinite  
CC proliferation of certain neoplastic diseases. The invention relates to  
CC antisense oligonucleotides directed against a nucleic acid sequence  
CC encoding either a chemokine receptor (CCR3), a common subunit of  
CC interleukin-4 (IL-4) and interleukin-13 (IL-13) receptors, or a common  
CC subunit of IL-4, IL-5 and GM-CSF receptors. The antisense  
CC oligonucleotides can be used in the treatment or prevention of asthma,  
CC allergy, hypersensitivity, general inflammation or cancer.

Sequence 19 BP; 2 A; 3 C; 8 G; 6 T; 0 other;

Query Match	89.5%;	Score 17;	DB 21;	Length 19;
Best Local Similarity	100.0%;	Pred. No. 91;		

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GTCGACGGGATGCT 19  
 |||  
 1 GTCGACGGGATGCT 17

Db

RESULT 16  
 ABX12694  
 ID ABX12694 standard; DNA; 19 BP.

XX  
 AC ABX12694;  
 XX  
 DT 10-MAY-2003 (first entry)  
 XX

DE Human IL-3/IL-5/GM-CSF receptor DNA, antisense oligonucleotide #7.

XX Human; inflammation; 2',6'-diaminopurine; DAP; antisense therapy;  
 KW DAP-modified oligonucleotide; pulmonary disease; respiratory disease;  
 KW neurological disease; cardiovascular disease; rheumatological disease;  
 KW digestive disease; cutaneous disease; ophthalmological disease;  
 KW urinary system disease; pathogen infection; genetic disease; cancer;  
 KW airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;  
 KW cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;  
 KW eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;  
 KW hyperesinophilia; cardiant; ophthalmological; cytostatic;  
 KW antiasthmatic; anti-allergic; antiinflammatory; immunosuppressive;  
 KW atopic disease; neoplastic cell proliferation; antisense; IL-3; IL-5;  
 KW interleukin-3 receptor; interleukin-5 receptor; GM-CSF; ss.

XX Homo sapiens.  
 OS  
 XX  
 XX MO2003004511-A2.  
 XX  
 PD 16-JAN-2003.  
 XX  
 PF 08-JUL-2002; 2002WO-CA01046.  
 XX  
 PR 06-JUL-2001; 2001US-303071P.  
 XX  
 XX (TOP1-) TOPIGEN PHARM INC.  
 XX  
 PI Renzi P, Allam M, Allakherdi Z;  
 XX  
 DR WPI: 2003-247944/24.  
 XX

PT Increasing in vivo efficacy of a nucleic acid molecule that is  
 PT administered to a mammal for inhibiting inflammation in mammals,  
 PT involves incorporating into the nucleic acid molecule at least one  
 PT nucleotide substitute -

PS Claim 28; Page 11; 63pp; English.

XX The present invention relates to a method for increasing the in vivo  
 CC efficacy of oligonucleotides and inhibiting inflammation. The  
 CC oligonucleotides comprise at least one nucleotide substitute of  
 CC 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide  
 CC substitutions are useful for increasing in vivo efficacy of a nucleic  
 CC acid molecule that is administered to a mammal. The DAP-modified  
 CC oligonucleotides are useful in antisense therapy for treating and/or  
 CC preventing pulmonary/respiratory diseases, neurological diseases,  
 CC cardiovascular diseases, rheumatological diseases, digestive diseases,  
 CC cutaneous diseases, ophthalmological diseases, urinary system diseases,  
 CC pathogen infections, genetic diseases, general inflammation and  
 CC cancer. The respiratory system disease is a sickness associated with  
 CC an inflammation of the lungs, the airways and/or the nose. The  
 CC respiratory system disease is selected from pulmonary fibrosis, adult  
 CC respiratory distress syndrome, cystic fibrosis, chronic obstructive  
 CC lung disease, chronic bronchitis, eosinophilic bronchitis, asthma,  
 CC allergy, allergic rhinitis, sinusitis and hyperesinophilia. The  
 CC DAP-modified oligonucleotides are more stable in the body, more  
 CC effective, and less toxic than standard antisense oligonucleotides.  
 CC DAP or its analogues are more effective than other substitutes of

CC adenosine. ABX12681-ABX12698 represent antisense oligonucleotides  
 CC for treating or preventing atopic diseases and neoplastic cell  
 CC proliferation.

XX

SQ Sequence 19 BP; 2 A; 3 C; 8 G; 6 T; 0 other.

Query Match 89.5%; Score 17; DB 25; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GTCGACGGGATGCT 19  
 |||  
 1 GTCGACGGGATGCT 17

Db

RESULT 17  
 ABX12701  
 ID ABX12701 standard; DNA; 19 BP.

XX  
 AC ABX12701;  
 XX  
 DT 10-MAY-2003 (first entry)  
 XX

DE Human IL-3/IL-5/GM-CSF receptor DNA, antisense oligonucleotide 107A-DAP.

XX Human; inflammation; 2',6'-diaminopurine; DAP; antisense therapy;  
 KW DAP-modified oligonucleotide; pulmonary disease; respiratory disease;  
 KW neurological disease; cardiovascular disease; rheumatological disease;  
 KW digestive disease; cutaneous disease; ophthalmological disease;  
 KW urinary system disease; pathogen infection; genetic disease; cancer;  
 KW airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;  
 KW cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;  
 KW eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;  
 KW hyperesinophilia; cardiant; ophthalmological; cytostatic;  
 KW antiasthmatic; anti-allergic; antiinflammatory; immunosuppressive;  
 KW atopic disease; neoplastic cell proliferation; antisense; IL-3; IL-5;  
 KW interleukin-3 receptor; interleukin-5 receptor; GM-CSF; ss.

XX Homo sapiens.  
 OS  
 XX  
 XX Synthetic.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH modified\_base 9 /\*tag= a  
 FT /mod\_base= OTHER  
 FT /note= "OTHER= DAP"  
 FT modified\_base 15 /\*tag= b  
 FT /mod\_base= OTHER  
 FT /note= "OTHER= DAP"

XX MO2003004511-A2.  
 XX  
 PD 16-JAN-2003.  
 XX  
 PF 08-JUL-2002; 2002WO-CA01046.  
 XX  
 PR 06-JUL-2001; 2001US-303071P.  
 XX  
 XX (TOP1-) TOPIGEN PHARM INC.  
 XX  
 PI Renzi P, Allam M, Allakherdi Z;  
 XX  
 DR WPI: 2003-247944/24.  
 XX

PT Increasing in vivo efficacy of a nucleic acid molecule that is  
 PT administered to a mammal for inhibiting inflammation in mammals,  
 PT involves incorporating into the nucleic acid molecule at least one  
 PT nucleotide substitute -

PS Examples; Page 18; 63pp; English.

XX The present invention relates to a method for increasing the in vivo

CC efficacy of oligonucleotides and inhibiting inflammation. The  
CC oligonucleotides comprise at least one nucleotide substitute of  
CC 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide  
CC substitutions are useful for increasing in vivo efficacy of a nucleic  
CC acid molecule that is administered to a mammal. The DAP-modified  
CC oligonucleotides are useful in antisense therapy for treating and/or  
CC preventing pulmonary/respiratory diseases, neurological diseases,  
CC cardiovascular diseases, rheumatological diseases, digestive diseases,  
CC cutaneous diseases, ophthalmological diseases, urinary system diseases,  
CC pathogen infections, genetic diseases, general inflammation and  
CC cancers. The respiratory system disease is a sickness associated with  
CC an inflammation of the lungs, the airways and/or the nose. The  
CC respiratory system disease is selected from pulmonary fibrosis, adult  
CC respiratory distress syndrome, cystic fibrosis, chronic obstructive  
CC lung disease, chronic bronchitis, eosinophilic bronchitis, asthma,  
CC allergy, allergic rhinitis, sinusitis and hyperesoinophilia. The  
CC DAP-modified oligonucleotides are more stable in the body, more  
CC effective, and less toxic than standard antisense oligonucleotides.  
CC DAP or its analogues are more effective than other substitutes of  
CC adenosine. The present sequence represents an antisense oligonucleotide  
CC used in the examples of the present invention.

CC  
XX Sequence 19 BP; 0 A; 3 C; 10 G; 4 T; 2 other;

Query Match 89.5%; Score 17; DB 25; Length 19;  
Best Local Similarity 89.5%; Pred. No. 91;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 19  
Db 1 GGGTCTGCAGCGGATGCT 19

RESULT 18  
ABX12702  
ID ABX12702 standard; DNA; 19 BP.

XX  
AC ABX12702;

XX  
DT 10-MAY-2003 (first entry)

XX  
DE Human IL-3/IL-5/GM-CSF receptor DNA, antisense oligonucleotide 107A-I.

XX  
KW Human; inflammation; 2',6'-diaminopurine; DAP; antisense therapy;  
KW DAP-modified oligonucleotide; pulmonary disease; respiratory disease;  
KW neurological disease; cardiovascular disease; rheumatological disease;  
KW digestive disease; cutaneous disease; ophthalmological disease;  
KW urinary system disease; pathogen infection; genetic disease; cancer;  
KW airway; nose; pulmonary fibrosis; adult respiratory distress syndrome;  
KW cystic fibrosis; chronic obstructive lung disease; chronic bronchitis;  
KW eosinophilic bronchitis; asthma; allergy; allergic rhinitis; sinusitis;  
KW hyperesoinophilia; cardiant; ophthalmological; cycostatic;  
KW antiaesthetic; anti-allergic; antiinflammatory; immunosuppressive;  
KW atopic disease; neoplastic cell proliferation; antisense; IL-3; IL-5;  
KW interleukin-3 receptor; interleukin-5 receptor; GM-CSF; ss.

XX  
OS Homo sapiens.  
XX  
OS Synthetic.

XX  
FH Key Location/Qualifiers  
FH modified\_base 9 /\*tag= a  
FT /\*tag= a  
FT /\*tag= i  
FT modified\_base 15 /mod\_base= i  
FT /\*tag= b  
FT /\*tag= i  
FT /mod\_base= i

XX  
PN MO2003004511-A2.

XX  
PD 16-JAN-2003.

XX  
PF 08-JUN-2002; 2002WO-CA01046.

XX

PR 06-JUN-2001; 2001US-303071P.

XX  
XX (TOP1-) TOPICEN PHARM INC.

XX  
PI Renzi P, Allam M, Allakhverdi Z;

XX  
DR WPI; 2003-247944/24.

XX  
PT Increasing in vivo efficacy of a nucleic acid molecule that is  
PT administered to a mammal for inhibiting inflammation in mammals,  
PT involves incorporating into the nucleic acid molecule at least one  
PT nucleotide substitute -

XX  
PS Examples; Page 18; 63pp; English.

XX  
CC The present invention relates to a method for increasing the in vivo  
CC efficacy of oligonucleotides and inhibiting inflammation. The  
CC oligonucleotides comprise at least one nucleotide substitute of  
CC 2',6'-diaminopurine (DAP) and/or its analogue. The DAP nucleotide  
CC substitutions are useful for increasing in vivo efficacy of a nucleic  
CC acid molecule that is administered to a mammal. The DAP-modified  
CC oligonucleotides are useful in antisense therapy for treating and/or  
CC preventing pulmonary/respiratory diseases, neurological diseases,  
CC cardiovascular diseases, rheumatological diseases, digestive diseases,  
CC cutaneous diseases, ophthalmological diseases, urinary system diseases,  
CC pathogen infections, genetic diseases, general inflammation and  
CC cancers. The respiratory system disease is a sickness associated with  
CC an inflammation of the lungs, the airways and/or the nose. The  
CC respiratory system disease is selected from pulmonary fibrosis, adult  
CC respiratory distress syndrome, cystic fibrosis, chronic obstructive  
CC lung disease, chronic bronchitis, eosinophilic bronchitis, asthma,  
CC allergy, allergic rhinitis, sinusitis and hyperesoinophilia. The  
CC DAP-modified oligonucleotides are more stable in the body, more  
CC effective, and less toxic than standard antisense oligonucleotides.  
CC DAP or its analogues are more effective than other substitutes of  
CC adenosine. The present sequence represents an antisense oligonucleotide  
CC used in the examples of the present invention.

XX  
SQ Sequence 19 BP; 0 A; 3 C; 10 G; 4 T; 2 other;

Query Match 89.5%; Score 17; DB 25; Length 19;  
Best Local Similarity 89.5%; Pred. No. 91;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGCT 19  
Db 1 GGGTCTGCAGCGGATGCT 19

RESULT 19  
AAK63218/C  
ID AAK63218 standard; cDNA; 256 BP.

XX  
AC AAK63218;

XX  
DT 06-NOV-2001 (first entry)

XX  
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8278.

XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cycostatic; gene therapy; vaccine; metastasis; ss.

XX  
OS Homo sapiens.

XX  
PN MO200157182-A2.

XX  
PD 09-AUG-2001.

XX  
PF 17-JAN-2001; 2001WO-US01354.

XX  
PR 31-JAN-2000; 2000US-0179065.

XX  
PR 04-FEB-2000; 2000US-0180628.

XX  
PR 24-FEB-2000; 2000US-0184664.



XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 256 BP; 67 A; 71 C; 71 G; 46 T; 1 other;  
Query Match 86.3%; Score 16.4; DB 22; Length 256;  
Best Local Similarity 94.4%; Pred. No. 2e+02; 1; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 GGGTCTGCAGCGGATG 18  
Db 207 GGGTCTGCAGCGGATG 190  
RESULT 20  
AA71002  
ID AAF71002 standard; DNA; 436 BP.  
XX  
AC AAF71002;  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE C. glutamicum SRT protein nucleotide sequence SEQ ID NO:37.  
XX  
KM Corynebacterium glutamicum; stress; resistance; tolerance; SRT;  
KM fine chemical production; organic acid; proteinogenic amino acid;  
KM nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
KM nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
KM carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;  
KM evolutionary study; environmental hazard; fermentation; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
XX WO200100804-A2.  
XX  
PD 04-JAN-2001.  
XX  
XX 23-JUN-2000; 2000WO-IB00922.  
XX  
XX 25-JUN-1999; 99US-0141031.  
XX 01-JUL-1999; 99DE-1030429.  
XX 01-JUL-1999; 99US-012692.  
XX 08-JUL-1999; 99DE-1031413.  
XX 08-JUL-1999; 99DE-1031457.  
XX 08-JUL-1999; 99DE-1031541.  
XX 09-JUL-1999; 99DE-1032209.  
XX 09-JUL-1999; 99DE-1032230.  
XX 14-JUL-1999; 99DE-1032914.  
XX 27-AUG-1999; 99DE-1040764.  
XX 27-AUG-1999; 99US-0151214.  
XX 31-AUG-1999; 99DE-1041382.  
XX  
PA (BADT ) BASF AG.  
XX  
PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Habernauer G, Lee H;  
XX Kim H;

DR WPI; 2001-061972/07.  
DR P-PSDB; AAB78889.  
XX  
PT New isolated Corynebacterium glutamicum nucleic acid encoding a stress,  
PT tolerance or resistance protein, for production or modulation of  
PT production of fine chemicals, such as, e.g. amino acids, lipids,  
PT carbohydrates, or enzymes -  
XX  
PS Claim 3; Page 172; 526pp; English.  
XX  
CC AAF70984 to AAF71133 encode the Corynebacterium glutamicum stress,  
CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.  
CC The C. glutamicum SRT genes (I) can be used in vectors (II) for  
CC expression in host cells and production of fine chemicals, such as, an  
CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),  
CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a  
CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic  
CC compound, a vitamin, a cofactor, a polypeptide, or an enzyme. The fine  
CC chemical production can be modulated. The presence of (I) or the SRT  
CC proteins (III) encoded by them are used for diagnosing the presence  
CC or activity of Corynebacterium diptheriae. (I), (II), (III) and host  
CC cells containing them can be used to map the genomes of organisms related  
CC to C. glutamicum, to identify and localise C. glutamicum sequences of  
CC interest, in evolutionary studies, in determination of SRT protein  
CC regions required for function, in modulating the SRT protein activity,  
CC and in modulating the activity of an SRT pathway. (II) are used to permit  
CC C. glutamicum to survive in an environment that is normally  
CC environmentally or chemically hazardous to it. (I) and protein molecules  
CC encoded by it increase the survival of C. glutamicum to chemical and  
CC environmental hazards and provide a means for continued growth and  
CC multiplication in large scale fermentative growth conditions. By  
CC increasing the growth rate or maintaining a normal growth rate in poor or  
CC toxic conditions, the yield, production and/or efficiency or production  
CC of fine chemicals from a culture may be increased.  
XX  
SQ Sequence 436 BP; 81 A; 102 C; 138 G; 115 T; 0 other;  
Query Match 86.3%; Score 16.4; DB 22; Length 436;  
Best Local Similarity 94.4%; Pred. No. 2.1e+02; 1; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 2 GGTCTGCAGCGGATG 19  
Db 401 GGTCTGCAGCGGATG 418  
RESULT 21  
AAK84431/C  
ID AAK84431 standard; DNA; 535 BP.  
XX  
XX AAK84431;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39243.  
XX  
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KM cytosolic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200157182-A2.  
XX  
XX 09-AUG-2001.  
XX  
PD 17-JAN-2001; 2001WO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225457.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249224.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249266.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259676.

(HUMA-) HUMAN GENOME SCI INC.  
XX  
PA  
XX  
Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
WPI; 2001-483426/52.  
XX  
XX  
DR  
XX  
XX  
PT  
XX  
XX  
XX  
PS  
XX  
XX  
CC  
CC  
CC  
CC  
activity, and can be used in gene therapy and vaccine production. (1)

AAK4951 to AAK64702 encode the human immune/hematopoietic antigen polypeptides,  
amino acid sequences given in AAK62170 to AAK91921. (1) have cytostatic  
activity, and can be used in gene therapy and vaccine production. (1)



CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169  
CC represent sequences used in the exemplification of the present invention.

SO Sequence 535 BP, 148 A, 159 C, 143 G, 85 T, 0 other;

Query Match 86.3%; Score 16.4; DB 22; Length 535;  
Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCTCTGACGCGGATGG 18  
|||  
Db 206 GGGCTCTGACGCGGATGG 189

RESULT 22  
AAK64430/C  
ID AAK64430 standard; DNA, 536 BP.  
XX  
AC AAK64430;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39242.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytoskeletal; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200157182-A2.  
XX  
PD 09-AUG-2001.  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230457.  
PR 06-SEP-2000; 2000US-0230458.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 13-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 DR WPI; 2001-483426/52.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX  
 PS Disclosure; SEQ ID NO 39242; 3071bp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAK62170 to AAK61921. (I) Have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK67694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169  
 CC represent sequences used in the exemplification of the present invention.  
 XX  
 XX Sequence 536 BP; 146 A; 161 C; 143 G; 86 T; 0 other;

Query Match 86.3%; Score 16.4; DB 22; Length 536;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGGTCTGCAGCGGATGC 18  
 |||||  
 Db 206 GGGTCTGCAGCGATGC 189  
 RESULT 23  
 AAS67706  
 ID AAS67706 standard; cDNA; 825 BP.  
 XX  
 AC AAS67706;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #3510.  
 XX  
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KM food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR P-PsDB; ABG03519.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 PS Claim 1; SEQ ID NO 3510; 103bp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 825 BP; 156 A; 278 C; 267 G; 124 T; 0 other;

Query Match 86.3%; Score 16.4; DB 23; Length 825;  
 Best Local Similarity 94.4%; Pred. No. 2.2e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 2 GGGTCTGCAGCGGATGCT 19

Db 665 GGCTGCAGCGGACGGT 682

## RESULT 24

AA583466 ID AA583466 standard; cDNA; 825 BP.

XX AC AA583466;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #19270.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB; ABG19279.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

XX Claim 1; SEQ ID No 19270; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AA564197-AA594564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIRO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 825 BP; 157 A; 285 C; 258 G; 125 T; 0 other;

Query Match 86.3%; Score 16.4; DB 23; Length 825;

Best Local Similarity 94.4%; Pred. No. 2.2e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GGCTGCAGCGGATGCT 19  
DB 668 GGCTGCAGCGGACGGT 685

## RESULT 25

AAA72267/C ID AAA72267 standard; cDNA; 1095 BP.

XX AC AAA72267;

DT 06-DEC-2000 (first entry)

DE Drosophila odorant receptor DORLU 1.1 cDNA.

KM Odorant receptor; Drosophila; olfactory receptor;

KW G protein-coupled receptor; GPCR superfamily; transgenic insect;

KW insect behaviour modification; pest control; pollinator attraction;

XX biosensor; odour detection; odour identification; apiculture; ss.

XX OS Drosophila melanogaster.

XX PN WO200043410-A2.

XX PD 27-JUL-2000.

XX PF 25-JAN-2000; 2000WO-US01823.

XX PR 25-JAN-1999; 99US-0117132.

XX (UYVA ) UNIV YALE.

XX Carlson JR, Kim J, Clyne PJ, Warr CG;

DR WPI: 2000-543246/49.

DR P-PSDB; AAB20932.

PT New nucleic acid encoding a Drosophila olfactory receptor, useful for  
PT identifying modulating agents -

XX Claim 3; Page 214-216; 303bp; English.  
XX Sequences AAA72236-A72284 represent cDNAs encoding Drosophila  
XX melanogaster odorant receptors (AAB20901-B20949). These proteins function  
XX as olfactory receptors, and are thought to be members of the G  
XX protein-coupled receptor (GPCR) superfamily, which is characterised by  
XX the presence of 7 transmembrane helices. Nucleic acids encoding the  
XX Drosophila odorant receptors may be used to generate expression  
XX constructs, host cells containing such constructs, and transgenic  
XX insects. Cells which express the odorant receptor genes may be used in  
XX methods to identify agents which modulate expression of these genes, and  
XX in methods to identify receptor binding partners. The Drosophila odorant  
XX receptor nucleic acids may also be used to identify corresponding genes  
XX in other insects, such as those which damage crops or transmit disease.

CC The odorant receptor proteins may be used to identify agents which  
CC modulate their activity, to identify binding partners, as antigens to  
CC raise antibodies, and in methods to modify insect behaviour. The proteins  
CC may be also be used in methods of behaviour modification. Such methods  
CC may be used to study or modify insect behaviour in response to odorants  
CC such as pheromones. Modification of insect behaviour has a wide range of  
CC applications, such as in pest control (e.g., by disrupting the feeding  
CC or mating behaviours of pest species), or for enhancing plant  
CC pollination (by attracting pollinator species). Odorant receptor proteins  
CC and/or nucleotides may also be used to identify appetite suppressants, to  
CC trap odours of a specific type, as biosensors for the detection of  
CC explosives, drugs, perfumes or pollutants, and in apiculture to modify  
CC the behaviour of bees, for example, to increase the production of royal  
CC jelly.

XX Sequence 1095 BP; 210 A; 336 C; 273 G; 276 T; 0 other;

Query Match 86.3%; Score 16.4; DB 21; Length 1095;

Best Local Similarity 94.4%; Pred. No. 2.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GGCTGCAGCGGATGCT 19  
DB 668 GGCTGCAGCGGACGGT 685

```
Db      |||||
      849 GGTCTGCAGCAGATGCT 832

RESULT 26
ID      AAA94826/c
XX      AAA94826 standard; DNA; 1194 BP.
XX
XX      AAA94826;
XX
XX      23-FEB-2001 (first entry)
XX
XX      Drosophila melanogaster odorant receptor DOR119 coding sequence.
XX
XX      Odorant receptor; fruit fly; DOR119; odour recognition; pest control; ds.
XX
XX      Drosophila melanogaster.
XX
XX      Key      Location/Qualifiers
XX      FT      1..1194
XX      CDS      /*tag= a
XX              /product= "DOR119"
XX              /partial
XX
XX      WO200050566-A2.
XX
XX      31-AUG-2000.
XX
XX      25-FEB-2000; 2000WO-US04995.
XX
XX      25-FEB-1999; 99US-0257706.
XX
XX      (UYCO ) UNIV COLUMBIA NEW YORK.
XX
XX      Vosshall LB, Amrein HO, Axel R;
XX
XX      WPI; 2000-572081/53.
XX
XX      P-PSDB; AAB26426.
XX
XX      Novel nucleic acid encoding an insect odorant receptor, for identifying
XX      PT      modulator compounds that are useful in controlling pest population
XX
XX      Disclousre; Page 58; 176pp; English.
XX
XX      The present sequence is the coding sequence for the previously identified
XX      CC      Drosophila melanogaster odorant receptor DOR119. The odorant genes and
XX      CC      proteins, such as those provided by the invention, are useful as they aid
XX      CC      in the study of the olfactory organ in mammals, as well as aiding the
XX      CC      understanding of the link between odour recognition and behaviour in
XX      CC      insects. They also enable the identification of compounds capable of
XX      CC      activating and inhibiting the receptors, allow the control of pest
XX      CC      populations via the use of alarm odour ligands and via the use of ligands
XX      CC      which interfere with the interaction between odorant ligands and
XX      CC      receptors associated with fertility.
XX
XX      SQ      Sequence 1194 BP; 237 A; 360 C; 307 G; 290 T; 0 other;

Query Match      86.3%; Score 16.4; DB 21; Length 1194;
Best Local Similarity 94.4%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
XX
XX      Drosophila melanogaster expressed polynucleotide SEQ ID NO 13559.
XX
XX      Drosophila; developmental biology; cell signalling; insecticide;
XX      KW      pharmaceutical; gene; ss.
XX
XX      Drosophila melanogaster.
XX
XX      WO200171042-A2.
XX
XX      27-SEP-2001.
XX
XX      23-MAR-2001; 2001WO-US09231.
XX
XX      23-MAR-2000; 2000US-191637P.
XX      PR      11-JUL-2000; 2000US-0614150.
XX
XX      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PWD, Myers EW;
XX
XX      WPI; 2001-656860/75.
XX      DR      P-PSDB; ABB62256.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      PT      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      PT      interactions -
XX
XX      Claim 1; SEQ ID NO 13559; 21pp + Sequence Listing; English.
XX
XX      The invention relates to an isolated nucleic acid detection reagent
XX      CC      capable of detecting 1000 or more genes from Drosophila. The invention is
XX      CC      useful in developmental biology and in elucidating cell signalling and
XX      CC      cell-cell interactions in higher eukaryotes for the development of
XX      CC      insecticides, therapeutics and pharmaceutical drugs. The invention
XX      CC      discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX      CC      sequences (AB101840-AB16175) and the encoded proteins
XX      CC      (ABB57737-ABB72072).
XX      CC      The sequence data for this patent did not form part of the printed
XX      CC      specification, but was obtained in electronic format directly from WIPO
XX      CC      at ftp.wipo.int/pub/published_pct_sequences.
XX
XX      SQ      Sequence 1197 BP; 239 A; 360 C; 307 G; 291 T; 0 other;

Query Match      86.3%; Score 16.4; DB 23; Length 1197;
Best Local Similarity 94.4%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

PR 12-JAN-2001; 2001NO-0000235.  
 PR 12-JAN-2001; 2001NO-0000239.  
 XX  
 PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.  
 PA (TIGR-) TIGR.  
 XX  
 PI Birkeland NK, Bidhammer I, Jonassen I, Jensen HB, Lien T;  
 PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;  
 PI Salzberg SL;  
 XX  
 DR WPI; 2002-557818/59.  
 XX  
 PT Novel DNA array useful for determining differential expression of  
 PT Methylococcus capsulatus genes, comprises polynucleotides or  
 PT oligonucleotides representative for a selective number of Methylococcus  
 PT capsulatus genes -  
 XX  
 PS Claim 19; Page 71-72; 678bp; English.  
 XX  
 CC The invention relates to a novel DNA array giving a representation of a  
 CC number of Methylococcus capsulatus genes. The method of the invention is  
 CC useful for determination of the differential expression of the genes of  
 CC M. capsulatus, and for studying gene expression on a genomic scale and in  
 CC gene expression assays of M. capsulatus genes. The sequences shown in  
 CC ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the  
 CC invention.  
 XX  
 SQ Sequence 1320 BP; 276 A; 409 C; 399 G; 236 T; 0 other;  
 XX  
 Query Match 86.3%; Score 16.4; DB 24; Length 1320;  
 Best Local Similarity 94.4%; Pred. No. 2.2e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 GGTCTGCAGCGGATGTT 19  
 Db 876 GGCTTGCAGCGGATGTT 859  
 XX  
 RESULT 29.  
 ABQ90381  
 ID ABQ90381 standard; DNA; 1443 BP.  
 XX  
 AC ABQ90381;  
 XX  
 DT 01-OCT-2002 (first entry)  
 XX  
 DE M. capsulatus gene #366 for DNA array.  
 XX  
 KM Micro array; gene; ds; differential expression; gene expression.  
 XX  
 OS Methylococcus capsulatus.  
 XX  
 PN WO200255655-A2.  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 14-JAN-2002; 2002WO-NO00019.  
 XX  
 PR 12-JAN-2001; 2001NO-0000235.  
 PR 12-JAN-2001; 2001NO-0000239.  
 XX  
 PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.  
 PA (TIGR-) TIGR.  
 XX  
 PI Birkeland NK, Bidhammer I, Jonassen I, Jensen HB, Lien T;  
 PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;  
 PI Salzberg SL;  
 XX  
 DR WPI; 2002-557818/59.  
 XX  
 PT Novel DNA array useful for determining differential expression of  
 PT Methylococcus capsulatus genes, comprises polynucleotides or  
 PT oligonucleotides representative for a selective number of Methylococcus

PT capsulatus genes -  
 XX  
 PS Claim 19; Page 224-225; 678bp; English.  
 XX  
 CC The invention relates to a novel DNA array giving a representation of a  
 CC number of Methylococcus capsulatus genes. The method of the invention is  
 CC useful for determination of the differential expression of the genes of  
 CC M. capsulatus, and for studying gene expression on a genomic scale and in  
 CC gene expression assays of M. capsulatus genes. The sequences shown in  
 CC ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the  
 CC invention.  
 XX  
 SQ Sequence 1443 BP; 259 A; 439 C; 447 G; 298 T; 0 other;  
 XX  
 Query Match 86.3%; Score 16.4; DB 24; Length 1443;  
 Best Local Similarity 94.4%; Pred. No. 2.3e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 GGTCTGCAGCGGATGTT 19  
 Db 496 GGCTTGCAGCGGATGTT 513  
 XX  
 RESULT 30  
 AAC79967/c  
 ID AAC79967 standard; cDNA; 1537 BP.  
 XX  
 AC AAC79967;  
 XX  
 DT 12-FEB-2001 (first entry)  
 XX  
 DE Human secreted protein encoding cDNA for gene 20.  
 XX  
 KM Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;  
 KM antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
 KM nootropic; neuroprotective; antibacterial; virucide; fungicide;  
 KM ophthalmological; vulnary; gene therapy; treatment; autoimmune disease;  
 KM hyperproliferative disorder; cardiovascular disorder; ocular disorder;  
 KM cerebrovascular disorder; nervous system disorder; infection; skin aging;  
 KM wound healing; epithelial cell proliferation; transplantation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058357-A1.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 23-MAR-2000; 2000WO-US07723.  
 XX  
 PR 26-MAR-1999; 99US-0126506.  
 PR 07-JAN-2000; 2000US-0174852.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX  
 DR WPI; 2000-611704/58.  
 DR P-PSDB; AAB45044.  
 XX  
 PT Nucleic acid molecules encoding human secreted proteins, used in  
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
 PT Parkinson's diseases and cancers -  
 XX  
 PS Claim 1a; Page 351-352; 418bp; English.  
 XX  
 CC This invention describes novel isolated nucleic acid molecules (I)  
 CC encoding a human secreted proteins (II) which have immunosuppressive,  
 CC antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiant,  
 CC vasotropic, cerebroprotective, nootropic, neuroprotective, antibacterial,  
 CC virucide, fungicide, ophthalmological and vulnary activity and can be  
 CC used for gene therapy. (I) and (II) are used to prevent, treat or  
 CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,  
 CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in

diagnosing a pathological condition or susceptibility to a pathological condition. The antibodies to (ii) can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities.

Query Match 86.3%; Score 16.4; DB 21; Length 1537;  
Best Local Similarity 94.4%; Pred. No. 2.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGTCTGCAGCGGATGG 18  
Db 1118 GGGTCTGCAGCGGATGG 1101

RESULT 31  
AAH65074  
ID AAH65074 standard; DNA; 1824 BP.  
AC AAH65074;  
DT 26-SEP-2001 (first entry)  
XX C glutamicum coding sequence fragment SEQ ID NO: 109.  
DE  
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;  
KM organic acid synthesis; ds.  
XX Corynebacterium glutamicum.  
OS  
XX EP1108790-A2.  
PN 20-JUN-2001.  
XX 18-DEC-2000; 2000EP-0127688.  
PF  
XX 16-DEC-1999; 99JP-0377484.  
PR 07-APR-2000; 2000JP-0159162.  
PR 03-AUG-2000; 2000JP-0280988.  
XX  
XX (KYOWA) KYOWA HAKKO KOGYO KK.  
PA  
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
PI  
XX MPI; 2001-376931/40.  
DR P-PSDB; AAG89855.  
XX  
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying mutation point of a gene, measuring expression of a gene, analysing expression profile or pattern of a gene and identifying homologous gene

Claim 8: SEQ ID NO: 109; 246bp + Sequence Listing; English.

The present invention provides a number of nucleotide and protein sequences from the Corynebacterium glutamicum. These are useful for identifying the mutation point of a gene derived from a mutant of Corynebacterium glutamicum, measuring expression amount and analysing the expression profile or expression pattern of a gene derived

from Corynebacterium glutamicum, and identifying a homologue of a gene derived from Corynebacterium glutamicum. Corynebacterium bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a nucleic acid described in the exemplification of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office.

Sequence 1824 BP; 375 A; 466 C; 538 G; 445 T; 0 other;

Qy 2 GGGTGCAGCGGATGG 19  
Db 301 GGGTGCAGCGGATGG 318

RESULT 32  
AAE71001  
ID AAE71001 standard; DNA; 1947 BP.  
AC AAE71001;  
DT 30-APR-2001 (first entry)  
XX  
XX C. glutamicum SRT protein nucleotide sequence SEQ ID NO:35.  
DE  
XX Corynebacterium glutamicum; stress; resistance; tolerance; SRT;  
KM fine chemical production; organic acid; proteinogenic amino acid;  
KM nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
KM nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
KM carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;  
KM evolutionary study; environmental hazard; fermentation; ds.  
XX  
XX Corynebacterium glutamicum.  
OS  
XX WO200100804-A2.  
PN 04-JAN-2001.  
XX  
XX 23-JUN-2000; 2000WO-IB00922.  
PF  
XX 25-JUN-1999; 99US-0141031.  
PR 01-JUL-1999; 99DE-1030429.  
PR 01-JUL-1999; 99US-0142692.  
PR 08-JUL-1999; 99DE-1031413.  
PR 08-JUL-1999; 99DE-1031457.  
PR 08-JUL-1999; 99DE-1031541.  
PR 09-JUL-1999; 99DE-1032209.  
PR 09-JUL-1999; 99DE-1032230.  
PR 14-JUL-1999; 99DE-1032914.  
PR 27-AUG-1999; 99DE-1040764.  
PR 27-AUG-1999; 99US-0151214.  
PR 31-AUG-1999; 99DE-1041382.  
XX  
XX (BADI) BASF AG.  
PA  
XX Pompeius M, Kroeger B, Schroeder H, Zelder O, Habernauer G, Lee H;  
PI Kim H;  
PI  
XX MPI; 2001-061972/07.  
DR P-PSDB; AAB78888.  
XX  
XX New isolated Corynebacterium glutamicum nucleic acid encoding a stress, tolerance or resistance protein, for production or modulation of production of fine chemicals, such as, e.g. amino acids, lipids, carbohydrates, or enzymes -

Claim 3, Page 166-169; 526pp; English.

CC AAF70984 to AAF71133 encode the *Corynebacterium glutamicum* stress,  
 CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.  
 CC The C. glutamicum SRT genes (I) can be used in vectors (II) for  
 CC expression in host cells and production of fine chemicals, such as, an  
 CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),  
 CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a  
 CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic  
 CC compound, a vitamin, a cofactor, a polyketide, or an enzyme. The fine  
 CC chemical production can be modulated. The presence of (I) or the SRT  
 CC proteins (III) encoded by them are used for diagnosing the presence  
 CC or activity of *Corynebacterium diptheriae*. (I), (II), (III) and host  
 CC cells containing them can be used to map the genomes of organisms related  
 CC to C. glutamicum, to identify and localise C. glutamicum sequences of  
 CC interest, in evolutionary studies, in determination of SRT protein  
 CC regions required for function, in modulating the SRT protein activity  
 CC and in modulating the activity of an SRT pathway. (II) are used to permit  
 CC C. glutamicum to survive in an environment that is normally  
 CC environmentally or chemically hazardous to it. (I) and protein molecules  
 CC encoded by it increase the survival of C. glutamicum to chemical and  
 CC environmental hazards and provide a means for continued growth and  
 CC multiplication in large scale fermentative growth conditions. By  
 CC increasing the growth rate or maintaining a normal growth rate in poor or  
 CC toxic conditions, the yield, production and/or efficiency or production  
 CC of fine chemicals from a culture may be increased.

XX Sequence 1947 BP; 399 A; 502 C; 562 G; 484 T; 0 other;

Query Match 86.3%; Score 16.4; DB 22; Length 1947;

Best Local Similarity 94.4%; Pred. No. 2.3e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GGTCTGCAGCGGATGCT 19

Db 401 GGTCTGCAGCGGAAAGT 418

RESULT 33

ABL06358 ABL06358 standard; cDNA; 3309 BP.

XX ABL06358;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 13556.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB62255.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from *Drosophila* and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 13556; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (AB857737-AB872072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 3309 BP; 878 A; 743 C; 832 G; 856 T; 0 other;

Query Match 86.3%; Score 16.4; DB 23; Length 3309;

Best Local Similarity 94.4%; Pred. No. 2.4e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GGTCTGCAGCGGATGCT 19

Db 1319 GGTCTGCAGCAGATGCT 1336

RESULT 34

AAV49652/C AAV49652 standard; DNA; 5802 BP.

XX AAV49652;

XX 23-OCT-1998 (first entry)

XX Human SRCR protein DNA.

XX Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;

XX nervous system; medullo-blastoma; glioma; breast; detection;

XX autoantibody; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 107..5464

XX /\*tag= a

XX /product= SRCR protein"

XX /note= "Scavenger receptor cysteine-rich domain

XX containing protein"

XX WO9830687-A2.

XX 16-JUL-1998.

XX 09-JAN-1998; 98WO-DE00096.

XX 18-JUL-1997; 97DE-1030997.

XX 09-JAN-1997; 97DE-1000519.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Mollenhauer J, Pouscka A;

XX WPI; 1998-399136/34.

XX P-PSDB; AAW64591.

XX Proteins containing scavenger receptor, cysteine rich domain -

XX useful for diagnosis and treatment of tumours

XX Claim 4; Fig 2; 54pp; German.

CC This sequence encodes a human protein which contains a SRCR (scavenger  
 CC receptor, cysteine-rich) domain. The gene and encoded protein can be used  
 CC to diagnose or treat tumours, particularly of the nervous system  
 CC (medullo-blastoma or glioma) or breast. The DNA sequence and probes  
 CC derived from it, are used to identify genes that express SRCR-domain  
 CC containing proteins, to determine the form in which these proteins exist

CC and to assess the significance of individual forms on cellular  
CC properties. The protein can be used to detect the presence of  
CC autoantibodies and antibodies which regulate its expression.  
XX  
SQ Sequence 5802 BP, 1253 A, 1614 C, 1576 G, 1359 T, 0 other;  
Query Match 86.3%; Score 16.4; DB 19; Length 5802;  
Best Local Similarity 94.4%; Pred. No. 2.4e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 GGGTCTGCAGCGGATCG 18  
DB 5435 GGGTCTGCAGCGGATCG 5418  
RESULT 35  
AAS33414  
ID AAS33414 standard; DNA; 8045 BP.  
AC AAS33414;  
XX  
XX 04-DEC-2001 (first entry)  
DE DNA encoding human secreted protein, Seq ID No 697.  
XX  
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KM rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;  
KM cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KM cytosolic; Alzheimer's disease; Parkinson's disease; human; cancer;  
KM multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KM Gaucher's disease; neurological disease; cerebrovascular disorder;  
KM thrombosis; wound healing; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO200155326-A2.  
PN  
XX 02-AUG-2001.  
PD  
XX 17-JAN-2001; 2001WO-US01347.  
PF  
XX 31-JAN-2000; 2000US-0179065.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-451931/48.  
DR  
XX  
XX New nucleic acids and polypeptides, useful for diagnosing, preventing  
PT or treating medical conditions -  
XX  
XX Disclosure; SEQ ID No 697; 753pp; English.  
PS  
XX  
XX The invention relates to novel isolated nucleic acid molecules (I)  
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,  
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,  
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in  
CC the prevention, treatment and diagnosis of diseases associated with  
CC inappropriate expression of secreted proteins. (I) and complementary  
CC sequences may also be used as DNA probes in diagnostic assays (e.g.  
CC polymerase chain reactions (PCR)) to detect and quantitate the presence  
CC of similar nucleic acid sequences in samples, and so which patients may  
CC be in need of restorative therapy. (II) may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of the secreted  
CC proteins. The anti-(II) antibodies and antagonists may also be used to  
CC down regulate expression and activity of (II). The anti-(II) antibodies  
CC may also be used as diagnostic agents for detecting the presence of (II)  
CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The  
CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.

CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/  
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia,  
CC cerebrovascular disorders), infections caused by bacteria, viruses and  
CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),  
CC agonists, antagonists and antibodies can also be used to promote wound  
CC healing, maintain organs before transplantation, and support cell culture  
CC of primary tissues. AAS33413-AAS33486 represent human secreted protein  
CC coding sequences, PCR primers, and related sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO  
CC at: ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 8045 BP; 1257 A; 2365 C; 2854 G; 1569 T; 0 other;  
Query Match 86.3%; Score 16.4; DB 22; Length 8045;  
Best Local Similarity 94.4%; Pred. No. 2.5e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 GGGTCTGCAGCGGATCG 18  
DB 1232 GGGTCTGCAGCGGATCG 1249  
RESULT 36  
AAS33413  
ID AAS33413 standard; DNA; 8047 BP.  
AC AAS33413;  
XX  
XX 04-DEC-2001 (first entry)  
DE DNA encoding human secreted protein, Seq ID No 696.  
XX  
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KM rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;  
KM cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KM cytosolic; Alzheimer's disease; Parkinson's disease; human; cancer;  
KM multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KM Gaucher's disease; neurological disease; cerebrovascular disorder;  
KM thrombosis; wound healing; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO200155326-A2.  
PN  
XX 02-AUG-2001.  
PD  
XX 17-JAN-2001; 2001WO-US01347.  
PF  
XX 31-JAN-2000; 2000US-0179065.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-451931/48.  
DR  
XX  
XX New nucleic acids and polypeptides, useful for diagnosing, preventing  
PT or treating medical conditions -  
XX  
XX Disclosure; SEQ ID No 696; 753pp; English.  
PS  
XX  
XX The invention relates to novel isolated nucleic acid molecules (I)  
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,  
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,  
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in  
CC the prevention, treatment and diagnosis of diseases associated with  
CC inappropriate expression of secreted proteins. (I) and complementary  
CC sequences may also be used as DNA probes in diagnostic assays (e.g.  
CC polymerase chain reactions (PCR)) to detect and quantitate the presence  
CC of similar nucleic acid sequences in samples, and so which patients may



CC be in need of restorative therapy. (II) may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of the secreted  
CC proteins. The anti-(II) antibodies and antagonists may also be used to  
CC down regulate expression and activity of (II). The anti-(II) antibodies  
CC may also be used as diagnostic agents for detecting the presence of (II)  
CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The  
CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
CC melanoma), neoplasms of the breast or liver, Sezary syndrome and  
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/  
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia,  
CC angina and thrombosis), infections caused by bacteria, viruses and  
CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),  
CC agonists, antagonists and antibodies can also be used to promote wound  
CC healing, maintain organs before transplantation, and support cell culture  
CC of primary tissues. AA53043-AA53486 represent human secreted protein  
CC coding sequences. PCR primers, and related sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO  
CC at: ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 8047 BP; 1258 A; 2367 C; 2853 G; 1569 T; 0 other;

Query Match 86.3%; Score 16.4; DB 22; Length 8047;

Best Local Similarity 94.4%; Pred. No. 2.5e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGCTCTGCAGCGCGATCG 18

1232 GGCTCTGCAGCGCGATCG 1249

RESULT 37

AA559568 standard; DNA; 21719 BP.

AA559568;

13-FEB-2002 (first entry)

Propionibacterium acnes immunogenic protein encoding DNA #63.

SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

dermatological; osteopathic; neuroprotectant; ds.

Propionibacterium acnes.

WO200181581-A2.

01-NOV-2001.

20-APR-2001; 2001WO-US12865.

21-APR-2000; 2000US-199047P.

02-JUN-2000; 2000US-208841P.

07-JUL-2000; 2000US-216747P.

(CORI-) CORIAX CORP.

Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

L'maisonneuve J, Zhang Y, Jen S, Carter D;

WPI; 2001-616774/71.

PS Claim 1; SEQ ID No 63; 1069bp; English.

XX Sequences AA559506-AA559804 represent DNA molecules encoding

CC Propionibacterium acnes immunogenic polypeptides. The proteins and their

CC associated DNA sequences are used in the treatment, prevention and

CC diagnosis of medical conditions caused by P. acnes. The disorders include

CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and

CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved

CC in infections of bone, joints and the central nervous system, however it

CC is particularly involved in the inflammatory lesions associated with acne

CC vulgaris. A method for detecting the presence or absence of P. acnes in a

CC patient comprises contacting a sample with a binding agent that binds to

CC the proteins of the invention and determining the amount of bound protein

CC in the sample. The polypeptides may be used as antigens in the production

CC of antibodies specific for P. acnes proteins. These antibodies can be

CC used to downregulate expression and activity of P. acnes polypeptides and

CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by

CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the

CC polypeptides shown in AA55069-AA55304 and AA67565-AA67566.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 21719 BP; 4285 A; 6649 C; 6206 G; 4578 T; 1 other;

Query Match 86.3%; Score 16.4; DB 23; Length 21719;

Best Local Similarity 94.4%; Pred. No. 2.6e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GGCTCTGCAGCGCGATCGT 19

8274 GGCTCTGCAGCGCGATCGT 8291

RESULT 38

AAV49655/C standard; DNA; 28720 BP.

AAV49655;

23-OCT-1998 (first entry)

Human SC3 DNA.

Scavenger receptor cytochrome rich domain; SRCR; diagnosis; treatment;

nervous system; medullo-blastoma; glioma; breast; detection; SC3;

autoantibody; ss.

Homo sapiens.

WO9830687-A2.

16-JUL-1998.

09-JAN-1998; 98WO-DE00096.

18-JUL-1997; 97DE-1030997.

09-JAN-1997; 97DE-1000519.

(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

Mollenhauer J, Pouscka A;

WPI; 1998-399136/34.

Proteins containing scavenger receptor, cytochrome rich domain -

useful for diagnosis and treatment of tumours

Disclosure; Fig 5c; 54pp; German.

This nucleotide sequence contains a fragment of a gene which is capable

of encoding a protein which contains a SRCR (scavenger receptor).



Wed Jan 14 07:58:27 2004

us-09-719-737-9.rng

Page 25

Search completed: January 13, 2004, 18:02:39  
Job time : 125.5 secs

---

